Markus Kaila

Safety and effectiveness of responsiveness monitoring in early intensive care

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo, July 29, 2015

Thesis supervisor:

Prof. Lauri Parkkonen

Thesis advisors:

M.Sc. (Tech.) Lasse Kamppari

Dr.Sc. (Tech.) Mika Särkelä

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Markus Kaila

Title: Safety and effectiveness of responsiveness monitoring in early intensive care

Date: July 29, 2015 Language: English Number of pages: 11+72

Department of Biomedical Engineering and Computational Science

Professorship: Medical Imaging

Supervisor: Prof. Lauri Parkkonen

Advisors: M.Sc. (Tech.) Lasse Kamppari, Dr.Sc. (Tech.) Mika Särkelä

Suboptimal sedation (over- and undersedation) of intensive care patients has been associated with adverse patient outcomes and in prolonged stay in hospital, leading to elevated treatment costs.

Several depth-of-anesthesia monitors have been tested for assessing the level of sedation. However, their safety and effectiveness has not been validated for regulatory authorities. To address the need for improved sedation monitoring, a novel measurement parameter called Responsiveness Index (RI) was developed by GE Healthcare in collaboration with Edinburgh Critical Care Research Group. The parameter quantifies patient's level of responsiveness from frontal electromyogram signal. The use of RI as an adjunct to other sedation practices is assumed to guide the nurses to administer sedatives according to patient's needs.

In this study, patient data from a randomized controlled pilot trial were analyzed in order to compare RI-augmented sedation monitoring to the current practice. The compared variables were RI values, incidences of deep sedation, and patient outcomes. As a secondary analysis, two patient subgroups in which the effects of RI monitoring were assumed to be most influential were analyzed.

The sample size in this study was small, and statistically significant results were not discovered. However, a trend was found indicating that patients who received RI-augmented sedation monitoring were more responsive, reached high states of responsiveness faster, and spent less time mechanically ventilated during the intervention period. Additionally, the patients in the RI-augmented group were significantly more responsive and received less analgesics when only the less responsive patients at baseline were analyzed. The results are promising, but a multicenter trial is needed in order to validate the safety and effectiveness of RI.

Keywords: sedation monitoring, intensive care unit, responsiveness

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Markus Kaila

Työn nimi: Responsiivisuusmonitoroinnin turvallisuus ja tehokkuus aikaisen vaiheen tehohoidossa

Päivämäärä: July 29, 2015 Kieli: Englanti Sivumäärä: 11+72

Lääaketieteellisen tekniikan ja laskennallisen tieteen laitos

Professuuri: Lääketieteellinen kuvantaminen

Työn valvoja: Prof. Lauri Parkkonen

Työn ohjaaja: DI Lasse Kamppari, TkT Mika Särkelä

Tehohoitopotilaiden suboptimaalinen sedaatio (liiallinen ja riittämätön sedaatio) on yhteydessä huonoihin potilaiden hoitotuloksiin ja pitkittyneisiin sairaalajaksoihin johtaen lisääntyneisiin hoitokustannuksiin.
Useita anestesian syvyyttä mittaavia monitoreita on testattu teho-osastopotilaiden sedaatiotasojen seurannassa, mutta niiden turvallisuutta ja tehokkuutta ei ole asianmukaisesti validoitu sääntelyviranomaisille. Vastatakseen kehittyneemmän sedaatiomonitoroinnin tarpeeseen, GE Healthcare on kehittänyt yhteistyössä Edinburghin tehohoidon tutkimusryhmän kanssa Responsiveness Index (RI) -parametrin. RI kvantifioi potilaan responsiivisuustason otsan elektromyografiasignaalista. Oletuksena on, että RI yhdessä muiden sedaatiokäytäntöjen kanssa voisi auttaa hoitajia annostelemaan sedatiiveja potilaan tarpeiden mukaan.
Tässä työssä analysoitiin potilasdataa, joka oli kerätty satunnaisotetusta vertailukokeesta. Kokeessa verrattiin RI:llä täydennettyä sedaatiomonitorointia nykyiseen hoitokäytäntöön. Ryhmiä vertailtiin seuraavien muuttujien osalta: RI arvot, syvän sedaation esiintyvyys ja potilaiden hoitotulokset. Jälkianalyysissä tarkasteltiin kahta potilaiden alaryhmää, joissa RI-monitoroinnin vaikutuksen odotettiin olevan tehokkain.
Analyysin perusteella vaikutti siltä, että potilaat joita hoidettiin RI:n kanssa olivat responsiivisempia, tulivat responsiivisiksi nopeammin ja viettivät vähemmän aikaa mekaanisesti ventiloituina interventiojaksolla. Nämä tulokset eivät kuitenkaan olleet tilastollisesti merkittäviä johtuen pienestä otoskoosta. RI-monitoroidut potilaat olivat merkittävästi responsiivisempia ja saivat merkittävästi vähemmän kipulääkkeitä, kun vain lähtötilanteessa vähemmän responsiiviset potilaat analysoitiin. Tulokset ovat lupaavia, mutta usean sairaalan kattava tutkimus suuremmalla otoskoolla tarvitaan validoimaan RI-monitoroinnin turvallisuus ja tehokkuus.

Avainsanat: sedaatiomonitorointi, teho-osasto, responsiivisuus

Preface

This Thesis was carried out at and funded by GE Healthcare Finland Oy. It has been a great pleasure to work with innovations that can improve quality and efficiency of healthcare.

First of all, I would like to thank my supervisors - Mika Särkelä for being a great scientific mentor, and Lasse Kamppari for introducing me to the subject. I also thank my professor Lauri Parkkonen for his advices on scientific writing and professor Tim Walsh from University of Edinburgh for sharing his clinical expertise.

A special thanks goes to my family and friends for their support over the years. Finally, I want to thank the GE Youth Group for awesome coffee breaks and making the nights unforgettable.

Helsinki, 29.7.2015

Markus Kaila

Contents

Terms, symbols and abbreviations

Terms

Agitation	Psychological and physical restlessness
Amnesia	Loss of memory caused by disease or physical or psychological trauma
Analgesia	A relief from pain
Anesthesia	A temporary state of uncosciousness
Anesthetic	A drug used to induce anesthesia
Analgesic	A drug used to provide analgesia
Analgo-sedation	A strategy that treats pain and achieves sedation at the same time
Anterograde amnesia	A condition in which events that occurred after the onset of amnesia cannot be recalled and new memories cannot be formed
Anxiolysis	Sedation or hypnosis used to reduce anxiety, agitation or tension
Apnea	Transient cessation of respiration
Ataxia	A set of neurological disorders that affect co-ordination
Benzodiazepine	Any of a family of minor tranquilizers that act against anxiety and convulsions; and produce sedation and muscle relaxation
Bispectral Index	One of a number of technologies used to monitor the depth of anesthesia by an algorithmic analysis of the electroencephalogram
Brainstem	Posterior part of the brain joining structurally continuously with spinal cord which provides the main motor and sensory neural excitation to the face and neck
Bradycardia	Slow heart rate
Bronchoscopy	Test to diagnose lung disease and view the airway
Constipation	A condition in which bowel movements are infrequent or incomplete
Cortical	Involving or resulting from the action of brain's outer layer of neural tissue
Deep sedation	A state where the patient does not respond to verbal stimulation
Delirium	A more or less temporary disorder of the mental faculties, as in fevers, disturbances of consciousness, or intoxication, characterized by restlessness, excitement, delusions, hallucinations, etc.

Electromyogram	A graphic record of the electric currents associated with muscular action
Encephalitis	A virus causing brain inflammation
Encephalopathy	Brain disease or malfunction of the brain
Endotracheal tube	A tube inserted into the airways to provide mechanical ventilation
Enteral tube	A common name for feeding tube used for patients who cannot obtain nutrition by mouth
Entropy measurement	A device to measure hypnosis in anesthesia
Epileptic seizure	A condition caused by excessive or synchronous neuronal activity in the brain
Extubation	The removal of patient from mechanical ventilator
Frontal electromyography	Refers to electromyography which is measured from upper facial muscles
Glabellar tap	A tap on the glabella (space between eyebrows and above the nose) to test the glabellar tap reflex
Hemodynamic	The branch of physiology dealing with the forces involved in the circulation of the blood
Hypotension	Low blood pressure
Hypoxia	Unadequate oxygen supply
Immuno-compromised patient	Patient whose immune system ability to fight infectious disease is compromised or absent
Inhibitory neurotransmitter	A substance that carries a signal from one nerve cell to another and has an inhibitory effect
Inter-rater reliability	Degree of agreement among raters
Intracerebral	Situated or occuring within the brain
Intravenous	Method to administer drugs directly into a vein
Intubation	Procedure involving the insertion of the tube used in mechanical ventilation
Light sedation	A state in which patient responds to verbal stimulation
Meningitis	Inflamation of protective brain membranes
Multicenter trial	A clinical trial performed at more than one clinic
Myocardial infarction	Occurs when myocardial ischemia exceeds a critical treshold and may result in irreversible cell damage to the heart
Myocardial ischemia	A blockage in the coronary arteries causing decrease in blood flow to the heart muscle
Nasogastric tube	A tube inserted through the nose to stomach to provide nutrition support

Nausea	A feeling of unease and discomfort of stomach
Neuromuscular blocking agent	A drug which blocks the muscular functions
Non-parametric method	A statistical method which does not make assumptions about distribution
Opiate	Any of several synthetic compounds having effects similar to natural opium alkaloids and their derivatives
Oxygen desaturation	Decrease in blood oxygen concentration
Patient State Index	EEG based measure intended to monitor the state of the brain
Patient ventilator dyssynchrony	Occurs when a patient's spontaneous respiratory efforts are no longer in synchrony with the mechanical ventilator
Pivotal trial	A trial used to make specific claims about efficacy and safety of the device
Post-traumatic stress disorder	A mental health condition caused by traumatic event
Prone positioning	A body position, in which one lies chest down
Sedation	The calming of mental excitement or abatement of physiological function, especially by the administration of a drug
Sedative	A drug which is used to conduct sedation
Self-extubation	An unwanted removal of critical tubes performed by the patient
Status epilepticus	A state where epileptic seizure lasts more than 5 minutes or where the epileptic seizures occure with high frequency
Suboptimal sedation	Over- or undersedation
Tracheostomy	Providing breathing support with a tube through opening at the front of the neck
Urinary retention	Inability to fully empty the bladder
Vascular catheter	A tube inserted to patient's vein allowing easy access to the bloodstream

Symbols

χ^2	Chi-squared distribution
E	Expected number of events
g	Weighting function of RI
n	Number of samples
N	Number of filtered power value samples
O	Observed number of events
p	p-value
P_{F}	Filtered fEMG power value
S	Scaling function of RI
t	Time
X^2	Chi-square

Abbreviations

AIDS	Acquired immune defiency syndrome
APACHE	Acute Physiological and Chronic Health Evaluation
BIS	Bispectral Index Scale
CCI	Charlson Comorbidity Index
EEG	Electroencephalogram
EMG	Electromyogram
fEMG	Frontal electromyogram
FDA	US Food and Drug Administration
ICU	Intensive care unit
ICU-LOS	Length of stay in ICU
IHI	Institute for Healthcare Improvement
i.v.	Intravenous
LOS	Length of stay in hospital
NMBA	Neuromuscular blocking agent
PSI	Patient State Index
PTSD	Post-traumatic stress disorder
RI	Responsiveness Index
RASS	Richmond Agitation–Sedation Scale
RSS	Ramsay Sedation Scale
SOFA	Sequential Organ Failure Assessment
VAP	Ventilator-associated pneumonia

1 Introduction

1.1 The economic burden of intensive care

Intensive care units (ICUs) are designed to treat and monitor critically ill and unstable patients. The patients admitted to the ICUs are the sickest in the hospital. They typically have failures in several organ systems and may be unable to breathe on their own [1]. The primary principle of intensive care is to aggressively support the patient's organ systems, not to perform complex interventions or investigations [2]. For example, many patients admitted to an ICU require mechanical ventilation to assist or replace spontaneous breathing [3]. Patients are admitted to the ICU for multiple reasons, for instance to treat severe illness or to monitor the patient after surgery or an accident [4].

The demand for intensive care services has been increasing worldwide as a result of aging population and the availability of more complex care procedures [5]. The United States has the highest number of ICU beds in the world [6], and more than five million people are admitted to the ICU annually [7]. According to The Society of Critical Care Medicine, annual critical care medicine costs in the US rose between 2000 and 2005 from $56.6 to $81.7 billion (44%) and represented 13% of all hospital costs [7]. Yet the costs are still increasing as in 2013 the critical care accounted for 20% of all hospital costs and 1% of US gross national domestic product [8]. In order to reduce ICU costs, the attention has recently focused on assessing patients with prolonged length of stay (LOS) in the ICU as well as strategies to improve quality of treatment. A study by Yaseen Arabi and colleagues found that patients with prolonged ICU stay formed only 11% of patients, but utilized 45% of ICU days. The pressure to decrease LOS in ICUs is not only about reducing costs, but also to tackle the adverse effects caused by prolonged stay. Prolonged ICU stay can adversely have a negative impact on patient's health status, since it increases the risk of infection and complications, and possibly even mortality. Operationally, increased LOS impacts also upon ICU bed availability, and leads to longer waiting times. [9]

ICU patients are often agitated and require sedatives to calm from mental distress. One way to decrease lenght of stay in intensive care unit (ICU-LOS) is to pay attention to sedation monitoring protocols and practice. As reported by Hughes and colleagues, suboptimal sedation (over- and undersedation) can increase risk of complications leading to elevated treatment costs [10]. For example, oversedation has been associated to lead to prolonged time in the mechanical ventilator, which has been shown to increase the risk of infections, and therefore, raise the costs of treatment [11]. Although valid clinical tools exist for monitoring the depth of sedation, they often rely on subjective assessments and require stimulating the patient, and therefore, may disturb the sleeping pattern. Several technologies for monitoring sedation level with electroencephalogram (EEG) analysis exist, but these methods were developed primarily for depth of anesthesia monitoring, and their validity for sedation monitoring is uncertain.

1.2 Responsiveness monitoring

To address the unmet need for improved sedation monitoring, a collaborative research programme between GE Healthcare and Edinburgh Critical Care Research Group developed Responsiveness Index (RI) – a novel parameter to quantify patient's level of responsiveness. The parametrs generates a responsiveness number (RI value), with a range of 0 (unresponsive) to 100 (highly responsive), from frontal electromyogram (fEMG) signal, which may help the caregiving nurse to adjust sedative medication according to the patient's needs. The use of Responsiveness Index as an adjunct to current clinical practices is expected to improve a range of patient-based and economic outcomes, including ICU-LOS.

1.3 Objectives of the study

This study looks to assess the safety and effectiveness of continuous Responsiveness Index-augmented sedation monitoring during early ICU care as a nurse decision-support tool. Data were analyzed from a randomized controlled pilot trial comparing sedation management using a protocol based on responsiveness monitoring against the current practice.

The main objective of this study was to test whether monitoring the patient with RI during early ICU care will decrease the period spent less responsive (RI < 20) without excess adverse events. The secondary objectives were to investigate whether monitoring the responsiveness of the patient reduces the incidences of deep sedation (a state where the patient does not respond to verbal stimulation); reduces the time to reach RI ≥ 20 and light sedation (a state where patient responds to verbal stimulation); and improves a range of patient-based outcomes.

In the primary analysis, all the patients recruited to the study were analyzed. In the secondary analysis, two patient subgroups were studied in which the effects of RI monitoring were expected to be most influental. The first subgroup consisted of patients with RI < 20 at the start of monitoring and the second subgroup consisted of patients who were deeply sedated at the start of monitoring. The objective of the secondary analysis was to test whether or not the patients in these subgroups were more affected by the study protocol. As this study was a pilot trial, the secondary analysis was performed in view of inclusion criteria for a pivotal trial. A pivotal trial is typically a large multicenter trial, which is used as the primary proof of safety and effectiveness of the device for regulatory authorities [12, 13].

This thesis is a part of a larger effort to provide clinical evidence to support the regulatory clearance from the US Food and Drug Administration (FDA). Due to the reason that the data were gathered from a pilot trial, and therefore, the sample size was small, statistically significant results were not expected to be found. However, rather than looking for statistically significant results between the compared variables, the aim was to find trends and provide descriptive statistics which could indicate that the tested protocol is feasible for a pivotal trial. By using the results as a reference, recommendations for the pivotal trial are given in Chapter 6.4.

2 Clinical background

During the past years, critical care researchers and physicians have started to recognize the adverse effects caused by suboptimal sedation. The recent publications in the field of critical care medicine suggest that agitation, as well as delirium and pain, should be evaluated systematically with evidence-based strategies [14, 15]. Although valid tools and strategies exist for sedation management, they are not widely used [16, 17].

This chapter will present the clinical background behind sedation and analgesia, the adverse effects caused by suboptimal sedation, and the current sedation management practices with their limitations. In addition, the severity-of-illness scores, which are used in this study to compare baseline differences, are briefly introduced at the end of this chapter.

2.1 Sedation and analgesia in the ICU

Critically ill patients – especially those receiving mechanical ventilation and other life-supporting interventions – are treated with sedatives as a standard procedure to reduce agitation [18]. Sedation is administered to reduce patients' discomfort and agitation during painful or distressing procedures that are frequently performed during ICU treatment. Sedation also helps the patient to cope with the emotional effects caused by critical illness experience. Agitation has been shown to be a commonly existing emotional state among ICU patients. Rowe and Fletcher have demonstrated that it occurs at least once in 71% of medical-surgical ICU patients and this is notably true for mechanically ventilated patients, who make up a major proportion of all ICU treated cases [19].

Agitation causes sympathetic stress response. If left untreated, the stress response can lead to negative prolonged acute consequences [10]. Most common symptoms caused by untreated stress response are difficulty of breathing, patient-ventilator dyssynchrony, elevated blood pressure and heart rate, combative behavior and post-traumatic stress disorder (PSTD). [20]

There are multiple factors causing agitation to the patient in an ICU setting. Commonly, agitation is a result of intractable or untreated pain, mechanical ventilation, invasive procedures, sleep deprivation, prone positioning, adverse drug effects, and alcohol and drug withdrawal [20]. In addition, unfamiliar environment, loss of self-control, inability to communicate, and isolation may have an effect in causing agitation [21]. Identifying the relationship between agitation and causative factors can be ambiguous, because they are often multifactorial (see Figure 1).

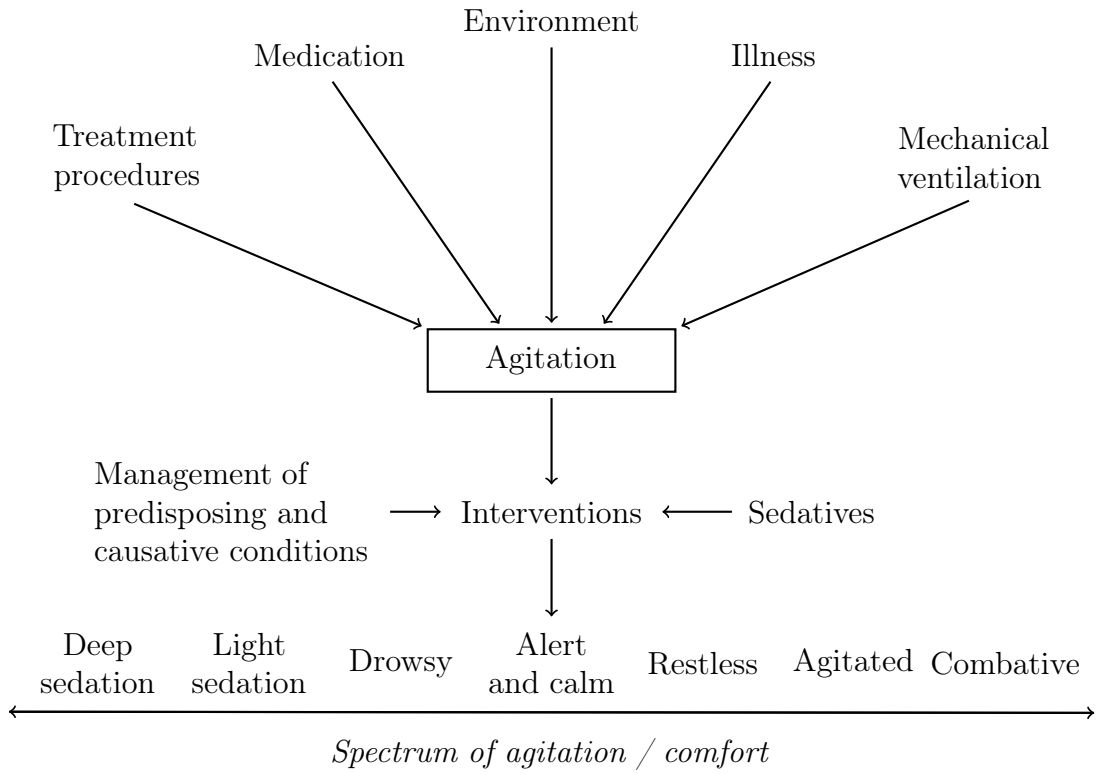


Figure 1: Causes of agitation. Modified from the original figure by Sessler and Muzevinch [3].

Agitation is inextricably linked to pain and delirium [22]. This means that pain and delirium can cause agitation and vice versa (see Figure 2). Additionally, treating agitation with doses of sedatives that alter consciousness may cause delirium [23].

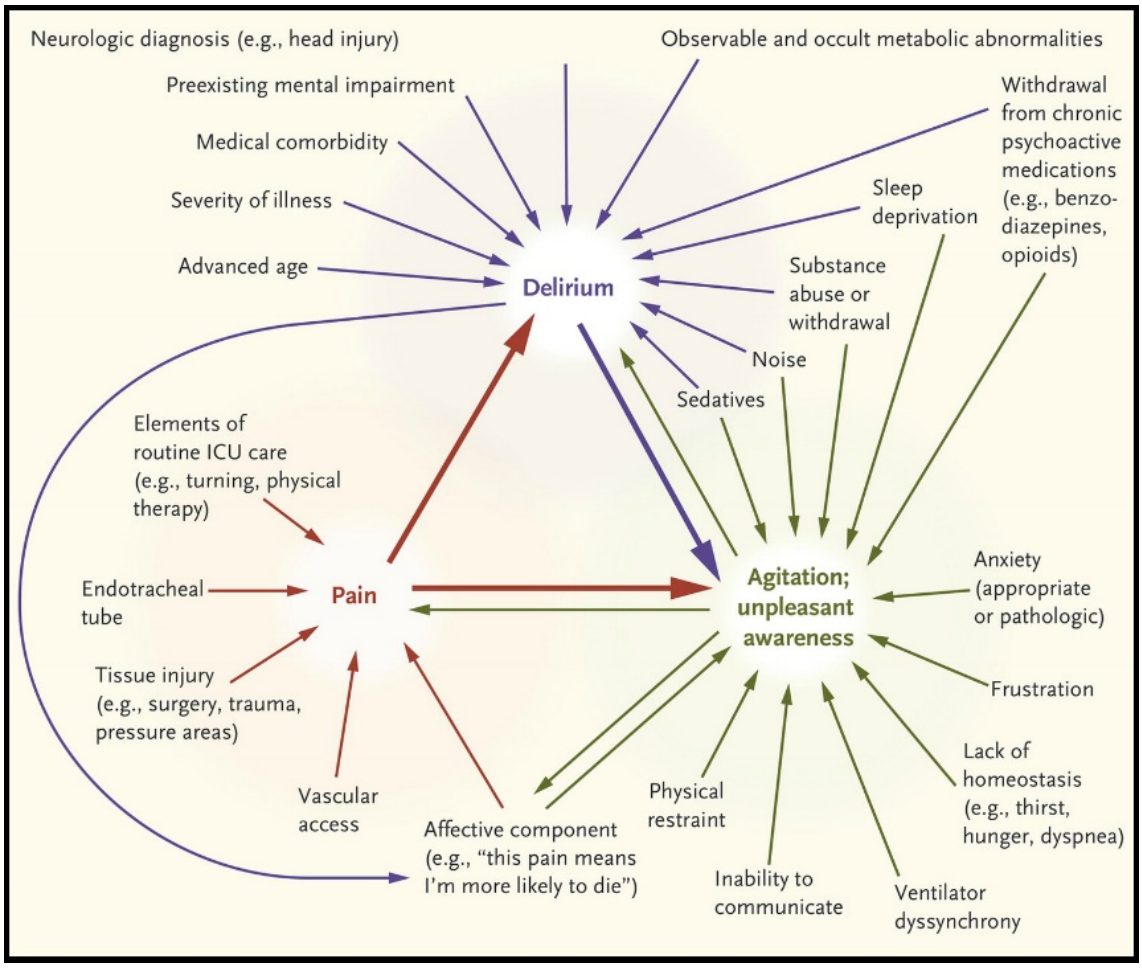


Figure 2: Connection between pain, agitation and delirium [22].

When pain contributes to development of agitation, agitation follows from the sensation of pain. Sedation aims to treat the anxiety component of pain, not the pain itself. [10] Attempts to control pain through only sedatives are inneffective and may lead to oversedation. Therefore, the first aim of analgo-sedation, a treatment focusing on pain and agitation simultaneously, is to ensure proper analgesia, and then control agitation [24]. Opiates, such as alfentanil and morphine, are the most frequently administered analgesics to treat pain in ICU setting. Even though they are effective in treatment of pain, opiates have significant side effects. They can cause respiratory depression, constipation, urinary retention, nausea, and confusion. [25] The balance between beneficial analgesia and adverse effects in opioid use is the limiting factor when conducting pain therapy [26].

Treating agitation with sedatives in a safe manner can be challenging. Sedatives have been categorized as "high-alert" drugs by the US based Institute for Healthcare Improvent (IHI). The category includes medicines which have been defined as "drugs that have the highest risk of causing injury when misused". [20, 27] Therefore, understanding the properties and limitations of sedatives is important when conducting sedative therapy. Bun and Dunn have described the ideal sedative as follows:

"*The ideal sedative has been described as one that works rapidly; provides anxiolysis, sedation, amnesia, or a combination of these; allows quick emergence when stopped; permits easy administration and adjustment of dosage; produces no active metabolities, significant adverse effects, or drug-drug interactions; and is inexpensive.*" They note that no sedative meets all these criteria and clinicians must therefore choose the right sedative according to the patient's condition [28]. Common and widely used sedatives are propofol and midazolam [19], which will be briefly presented in the following sections.

2.1.1 Propofol

Propofol is a sedative with hypnotic properties and it has a rapid onset and offset of action [29]. Because the concentration in blood drops quickly after the administration has stopped, it can be used in patients who need rapid awakening [28]. The rapid onset and offset allow greater control over the depth of sedation [30]. Propofol provides adequate sedation with a similar proportion to midazolam, but the recovery rate is faster with the patients for whom propofol is administrated [29].

Table 1: Therapeutic indications and possible adverse effects reported by the manufacturer for propofol [31].

Therapeutic indications	**Adverse effects**
Short-acting intravenous general anesthetic used for: • induction and maintenance of general anesthesia; • sedation for diagnostic and surgical procedures, alone or in combination with local or regional anesthesia; and • sedation of adult ventilated patients.	Common adverse events are: headache in the recovery phase; bradycardia; hypotension; transient apnea during induction; nausea and vomiting during recovery phase; and local pain on induction.

2.1.2 Midazolam

Midazolam is a benzodiazepine, which causes sedation and anterograde amnesia by affecting the inhibitory neurotransmitters. It is recommended for treatment of acute agitation due to its rapid onset (two to five minutes). Midazolam has been shown to accumulate and cause prolonged sedation in the ICU, hence it is not recommended for longer use than 24 hours [28]. A literature review conducted by Audrey Shafer on complications of sedation with midazolam in the ICU highlight that continuous infusion of midazolam is an effective way to sedate ICU patients, but it may cause complications. The reported complications were development of tolerance particulary in longer-term infusions, prolonged sedation due to accumulation of the drug, and cluster of symptoms that emerge during withdrawal from the drug. [32]

Table 2: Therapeutic indications and possible adverse effects reported by the manufacturer for midazolam [33].

Therapeutic indications	**Adverse effects**
For sedation by continuous infusion in patients in intensive care.	The following undesirable effects have been reported (very rarely) to occur when midazolam is injected: confusional state; euphoric mood; hallucinations; paradoxical reactions such as agitation; involuntary movements; prolonged sedation; decreased alertness; sleepiness; headache; dizziness; ataxia; cardiac arrest; respiratory depression; and anterograde amnesia.

2.2 Suboptimal sedation

As noted earlier in this Thesis, suboptimal sedation is associated with adverse events and worse clinical outcomes, leading to an increased ICU-LOS [10, 19].

Undersedation can lead to multiple undesirable effects, such as self-extubation or even combatitive behaviour against the care providers. It can occur in critically ill patients with as high rates as 75% of all patient during ICU stay. Most common responses to undersedation are self-removal of critical tubes and vascular catheters, which place the patient in danger from loss of treatment devices. [3] Mion and colleagues investigated the patient-initiated device removal in ICUs and found that self-removal of devices occured in 22 incidences out of 1000 patient-days [34]. In addition to device removal and combative behaviour, undersedation can lead to intense stress response [3], hence affect the patient comfort. In the worst-case scenario, the mental stress resulting from being consciouss through painful and terrifying procedures can lead to post-traumatic stress disorder [35].

Oversedation is a general problem in the ICU. A systematic review on studies of sedation practices conducted by Jackson et al. showed that substantial number of suboptimal sedation incidences were reported in all reviewed studies, with tendency to oversedation [36]. Oversedation can hamper the recovery by causing delirium and immobility, and therefore, lead to prolonged LOS. [20] Another important consequence of oversedation is the supression of spontaneous breathing reflex [37], which can delay the weaning from ventilator and thus, increase the duration of mechanical ventilation. In addition, oversedated patients are not able to influence their own treatment, as they are unable to interact with their family and environment.

Oversedating patients can lead to a vicious cycle (see Figure 3). The excessive sedation creates a dynamic loop that can aggravate muscle waisting, thus lead to a prolonged time in mechanical ventilator [38]. The additional time spent in mechanical ventilator has been shown to increase the risks of lung injury and infections [11], which result in unimproving health status of the patient, and therefore, may lead nurses to increase sedation infusion.

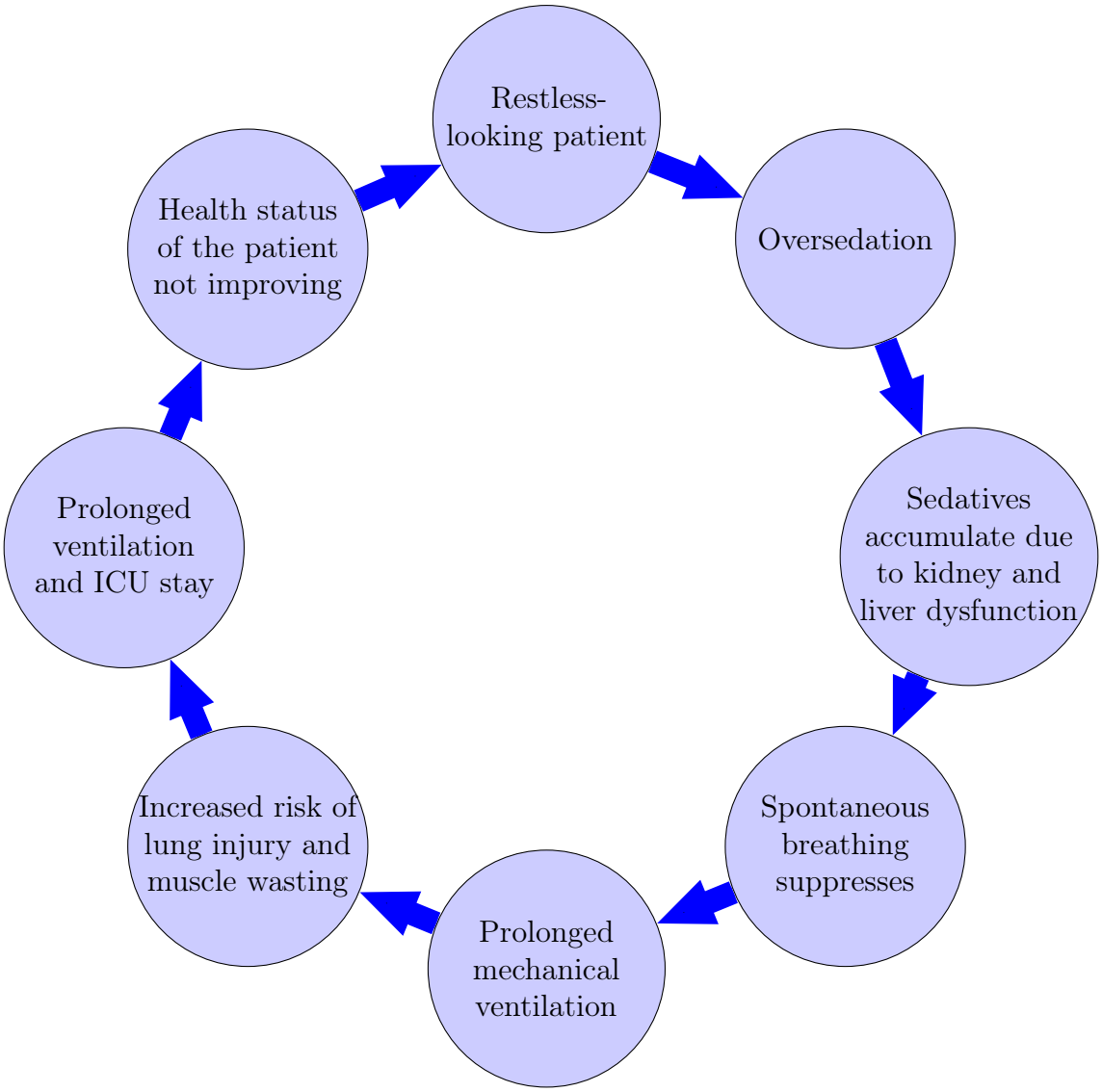


Figure 3: Vicious cycle of oversedation. Modified from Vasilevskis et al. [38].

Even though oversedation can have severe consequences on the patient's health, visually oversedated patients look peaceful and comfortable, and deeply sedated states can visually mimic peaceful sleep. Therefore, the care-giving nurses can perceive oversedation as kindness towards the patient. [39]

It must be noted that some conditions and situations may require deeper sedation for depression of consciousness to induce a state where the patient is not easily arousable, but still responds to painful stimulation [40]. Examples of conditions where deep sedation is required are ventilator dyssynchrony secondary to hypoxia, hyperactive airway disease infuriated by anxiety, and very uncomfortable procedures such as inserting the endotracheal tube or bronchoscopy. [20]

Achieving balanced sedation is particularly challenging in mechanically ventilated patients, as a review by Jackson and the research team highlighted that potentially 40–60% of mechanically ventilated patients are oversedated [36]. Being on a mechanical

ventilator can cause strong feeling of agitation, as the patient is often aware that he or she is in a life-threatening situation without control or means to communicate [41].

As oversedation prolongs the time spent in a mechanical ventilator, it can lead to ventilator-associated pneumonia (VAP). VAP is the most common hospital-acquired infection among ICU patients. Between 10% and 20% of patients who received more than 48 hours of mechanical ventilation will develop VAP. It may lead to significant morbidity, a two-fold mortality rate and elevated costs. In a study by Safdar and colleagues, the research team reported that VAP increased mean ICU-LOS for 6.1 days and added additional hospital costs over $10 000 per patient. [11]

2.3 Sedation management

Modern practical guidelines stress the importance of adopting patient-focused approach in sedation management, which aims to treat the patient with the lowest effective dose of drugs. Proper sedation management controls the following elements: monitoring the depth of sedation; treating the patient with correct sedative; assuring that the subject is kept on the lightest level of sedation as possible; and establishing goals of sedative therapy. [3] At present, sedation management is typically implemented through sedation protocols, which include clinical assessments linked to decision-making strategies, and sedation holds.

Adopting sedation management protocol has been proven to lead to better patient outcomes. In 2010, Scrobik and colleagues published results from an intervention study, in which the research team examined the effects of using a sedation management protocol. The data were collected from adult patients admitted to an ICU during two periods: before (610 patients) and after (604 patients) the implementation of the protocol. Patients who received protocol based care had a lower 30-day mortality risk than those who had received non-protocolized care. They also spent less time in mechanical ventilator, and medication induced coma rates were lower. [23]

As suboptimal sedation has a significant impact on patient outcomes, therefore the aim of sedation monitoring is to avoid incidences of both over- and undersedation. Sedation monitoring is mostly executed through clinical assessments which are linked to decision-making strategies to titrate the sedatives according to patient's needs. The clinical assessments are performed by observing the behaviour of the patient by using structured sedation scales. When applying these scales, the care provider gives different types of standard stimuli and assesses the patient's response. The arousal level tested typically ranges from alert to comatose. [3] The commonly used sedation scales are Richmond Agitation-Sedation Scale (RASS) and Ramsay Sedation Scale (RSS).

2.3.1 The Richmond Agitation-Sedation Scale

According to Sessler and colleagues, the Richmond Agitation-Sedation Scale (RASS) is the most valid and reliable sedation assessment tool for adult ICU patients to monitor the depth of sedation [3]. RASS was developed at Commonwealth University in Richmond, USA by a team of physicians, nurses and pharmacists [42]. It is a

ten-point scoring system based on the response to the assessor's voice and to physical stimulation. The range of the scale varies from –5 (unrousable) to +4 (combative). A score of 0 signifies that the patient is alert and calm. [43] RASS has the advantage of separating verbal from physical stimulation, allowing the patient's level of arousal to be graded on the basis of the stimulus [44].

In 2003, Ely and the research team discovered that RASS has excellent inter-rater reliability by analyzing 290 paired observations by nurses. Additionally they tested the face validity of RASS via survey of 26 critical care nurses, and the results showed that 92% agreed or strongly agreed with the RASS scoring system. The same survey demonstrated 81% of the nurses agreed or strongly agreed that RASS provided a consensus for goal-directed delivery of sedatives. [44]

Table 3: Richmond Agitation-Sedation Score [42].

Score	**Term**	**Description**
+4	Combative	Overly combative or violent, immediate danger to staff
+3	Very agitated	Pulls on or removes tubes or catheters or has aggressive behavior toward staff
+2	Agitated	Frequent nonpurposeful movement or patient ventilator dyssynchrony
+1	Restless	Anxious or apprehensive but movements not aggressive or vigorous
0	Alert and calm	
–1	Drowsy	Not fully alert, but has sustained (more than 10 seconds) awakening, with eye contact/eye opening to voice
–2	Light sedation	Briefly (less than 10 seconds) awakens with eye contact to voice
–3	Moderate sedation	Any movement (but no eye contact) to voice
–4	Deep sedation	No response to voice, but any movement to physical stimulation
–5	Unarousable	No response to voice or physical stimulation

The RASS scoring is done in accordance with the following procedure [42]:

1. Observe patient. Is patient alert and calm (score 0)?
2. Does patient have behavior that is consistent with restlessness or agitation (score +1 to +4 using the criteria listed above)?

3. If patient is not alert, in a loud speaking voice state patient's name and direct patient to open eyes and look at speaker. Repeat once if necessary.
   (a) Patient has eye opening and eye contact, which is sustained for more than 10 seconds (score –1).
   (b) Patient has eye opening and eye contact, but this is not sustained for 10 seconds (score –2).
   (c) Patient has any movement in response to voice, excluding eye contact (score –3).
4. If patient does not respond to voice, physically stimulate patient by shaking shoulder and then rubbing sternum if there is no response to shaking shoulder.
   (a) Patient has any movement to physical stimulation (score –4).
   (b) Patient has no response to voice or physical stimulation (score –5)

2.3.2 The Ramsay Sedation Scale

The Ramsay Sedation Scale (RSS) was the first widely utilized sedation scale for ICU patients [3]. It was first described by Ramsay et al. in 1947 and it was designed to test patient's rousability [26, 45]. Now it is one of the most widely used sedation scale [46]. The Ramsay Sedation Scale has six different levels measuring the patient's rousability from conscious and agitated state (level 1) to unresponsive (level 6). The stimulus was designed not to be painful, and it should not rouse a sleeping patient excessively to impede normal sleeping pattern. [26]

De Jonghe and colleagues have stated that RSS inhibits a satisfactory inter-rater reliability and high correlation with other sedation scales [47]. However, compared to RASS, RSS does not make a clear distinction between levels of agitation.

Table 4: The Ramsay Sedation Scale [26].

Level	**Characteristics**
1	Patient awake, agitated or restless
2	Patient awake, co-operative, orientated and tranquil
3	Patient drowsy, with response to commands
4	Patient asleep, brisk response to glabellar tap or loud auditory stimulus
5	Patient asleep, sluggish response to stimulus
6	Patient has no response to firm nailbed pressure or other noxious stimuli

The Ramsay Sedation Scoring is done by using the following steps: observation of behaviour (score 1 or 2), followed (if necessary) by assessment of response to voice (score 3), followed (if necessary) by assessment of response to loud auditory stimulus or light glabellar tap (score 4 to 5). [26]

2.3.3 Sedation holds

Sedatives can be administered by continuous infusion or as boluses when required. The most common way is to use continuous infusion, which provides a steady sedation with less episodes of occasional agitation. [19] However, continuous infusion may lead to prologed mechanical ventilation and ICU stay. In addition, the patient's neurological status cannot be accurately monitored. A study by Kollef and the research team revealed that the duration of mechanical ventilation was significantly longer for patients receiving continuous infusion of sedatives as opposed patients not receiving continuous infusion of sedatives. They suggested that strategies targeted at reducing the use of continuous infusion could decrease the duration of mechanical ventilation for some patients. [48]

One way to tackle the adverse effects caused by continuous infusion is to put the administration of sedatives on hold by decreasing the level of sedation slowly until patient is responsive. [49, 50] The sedation interruptions allow clinicians to find the optimal sedation level while maintaining patient comfort [49].

Multiple studies have demonstrated the benefits of using daily sedation holds as standard practice in the ICU. For example, in 2000, Kress et al. compared standard practice without sedation scoring or pre-defined protocols against using daily sedation holds in ICU patients. The authors found a significant decrease in duration of mechanical ventilation and ICU-LOS. Additionally, the patients receiving daily sedation holds did not have abnormally high rate of adverse events. [49]

In a study conducted by Girard and the research team in 2008, sedation holds were shown to significantly decrease in duration of mechanical ventilation, ICU-LOS and hospital LOS. Additionally, they found decrease in mortality rate after one year of the study. [51]

2.3.4 Limitations of current sedation practices

Protocolized sedation management, including sedation scales and daily interruption of sedatives, have been proven to lead to better patient outcomes. These protocols have been demonstrated to shorten the ICU lenght of stay, reduce the time spent in mechanical ventilator, and minimize the effects of sedative accumulation. [52]

Even though the sedation protocols have been demonstrated effective, they are not widely used. A survey study conducted by Tanios and the research group showed that the daily sedation interruption was used only by 40% of critical care clinicians and 36% did not have sedation management protocol in use. Protocols were not followed for the following reasons: no physician order (35%), lack of nursing support (11%), and fear of oversedation (7%). [16]

Sedation guidelines and protocols may be ignored unintentionally by the physicians. In 2000, Slomka and colleagues investigated the clinicians values and perceptions on use of clinical practice guidelines for sedation. The researchers came to a conclusion that physicians may think that they are following sedation guidelines when they are not. The research showed that 69% of physicians reported following guidelines, but in reality their actual adherence was only 20%. [17]

The problem with clinical scoring systems is that they are subjective, and thus, can be subjected to variability between interpreters. According to Weinert and colleagues, the subjective assessments are affected by social, personal, and professional factors [53]. Additionally, they do not discriminate deeper levels of sedation clearly, and require stimulating the patient, hence can cause sleep disruption and distress [54].

In addition, the choise of sedation management protocol can also be challenging. There is no universally accepted sedation management protocol and selecting the appropriate one can depend on multiple factors. Adopting a certain sedation management protocol can be affected by regional preferences, patient history, institutional bias and patient and clinician variability [10].

2.4 Severity-of-illness scores

Severity-of-illness scores have been designed compare the ICU patients' health status. These scores are calculated from physiological variables measured typically within the first 24 hours of ICU stay. In research, the severity of illness scores are often used for risk adjustment in ICU outcome studies [55]. They can also be used for quality assurance purposes in randomized controlled trials to determine whether the compared groups are similar in terms of therapy regimen [56]. Commonly used severity-of-illness scores are Acute physiological and chronic health evaluation (APACHE II), The Sequential Organ Failure Assessment (SOFA) score, and Charlson Comorbidity Index (CCI), which are briefly introduced in the following subchapters.

2.4.1 Acute physiological and chronic health evaluation

Acute physiological and chronic health evaluation (APACHE II) is a severity-of-illness classification system, which is used in the ICU to help to estimate the prognosis of the patient. [57] In APACHE II, there are 12 physiological variables and the effects of age and chronic health status are incorporated directly into the model. The worst recorded value, for every physiological variable during first 24 hours after ICU admission together with the weighted chronic health status and age, are used to form a single increasing score with a range from 0 to 71. [58]

In APACHE II scoring system, the patient receives points depending on the difference between the measured physiological value and normal range defined by the system. It is performed during the first 24 hours after the ICU admission. If the measured physiological value is in the normal range, the patient will receive 0 points, while the maximum points for high abnormality are +4. The investigated variables are listed in table 5. [57]

Table 5: APACHE II physiological variables with definitions/clarifications. Modified from a publication by Knauss et al. [57].

Physiological variable	**Definition or clarification**
Body temperature	Measured from rectal in celcius
Arterial pressure	Average blood pressure
Heart rate	Ventricular response
Respiratory rate	Non ventilated or ventilated
Oxygenation	Concentration of oxygen in blood
Arterial pH	Acidity of arterial blood
Sodium (serum)	Amount of sodium in blood serum
Potassium (serum)	Amount of potassium in blood serum
Creatinine (serum)	Amount of creatine in blood (patients receiving chronic dialysis get double points)
Hematocrit	Volume percentage of red blood cells in blood
White blood count	Number of white blood cells in blood
Glasgow coma score	A neurological scale recording the conscious state of patient

Chronic health points are given if the patient has a history of severe organ insufficiency or is immuno-compromised. The points are given as follows: two points for elective postoperative patient with immuno-compromise or history of severe organ insufficiency, and five points for nonoperative patient or emergency postoperative patient with immuno-compromise or severe organ insufficiency. [57] Patients will also receive points depending on age according to Table 6.

Table 6: Points given based on age [57].

Age (yrs)	**Points**
$\leq$44	0
45–54	2
55–64	3
65–74	5
$\geq$75	6

2.4.2 Sequential Organ Failure Assessment score

The Sequantal Organ Failure Assessment (SOFA) scoring is used to analyze patient's organ dysfunction or failure, and it can be used to evaluate morbidity and predict mortality [59]. It measures the function of six major organ systems; cardiovascular, respiratory, renal, hepatic, central nervous system, and coagulation [60]. Each system gets a score from 0 (normal) to 4 (most dysfunction), thus the maximum score is 24 [61].

2.4.3 Charlson Comorbidity Index

The Charlson Comorbidity Index uses a total of 22 common conditions (see Table 7) to predict mortality of the patient. The score of each condition has been weigted by the relative risk of mortality associated with each disease. The scores for each disease are summed together to form the index, with a range from 0 (lowest risk of mortality) to 37 (highest risk of mortality). [62]

Table 7: Charlson Comorbidity Index variables. Modified from the publication by Kastner et al. [63].

Condition	**Definition**	**Score**
Myocardial infraction	Occurs when blood flow stops to a part of heart muscle due to a blockage	1
Congestive heart failure	A condition in which heart muscle does not pump blood correctly	1
Peripheral vascular disease	Blood vessel disease located outside the heart and the brain	1
Cerebrovascular disease	A condition caused by inadequate supply of blood to the brain	1
Dementia	Set of symptoms including memory loss and gradual decrease in ability to think	1
Chronic pulmonary disease	A set of lung diseases which make the breathing difficult	1
Connective tissue disease	A disease that attacks the core components of connective tissue	1
Peptic ulcer disease	A break in the lining of stomach or first part of the duodenum	1
Mild liver disease	Patients with cirrhosis without portal hypertension or chronic hepatitis (inflammation of the liver)	1
Diabetes	A condition which causes the blood sugar levels to become too high for over a prolonged period	1
Hemiplegia	Paralysis of one side of the body	2
Moderate or sever renal disease	Moderate renal insufficiency includes patients with high serum creatine levels: severe renal disease includes patients on dialysis, those who had a transplant, and those with renal failure	2
Diabetes with or end organ damage	Diabetics whose retina or nerves have damaged	2
Any tumor	Patients with solid tumors without spread of cancer	2
Leukemia	A cancer of the blood cells	2
Lymphoma	A cancer that affects the immune system	2
Moderate or severe liver disease	Patients with chirrosis and hypotension	3
Metastatic solid tumor	A solid tumor which has spread to other parts of the body	6
AIDS	Acquired immune defiency syndrome	6

3 Technological background

In order to have a regulatory approval for a medical device from the US Food and Drug Administration, the manufacturer must demonstrate valid scientific evidence that provides reasonable assurance that the device is safe and effective for its intended use [64]. The safety and effectiveness is usually demonstrated through pivotal trials. The pivotal trials are typically large multicenter trials [65] conducted after the study protocol has been tested with a small sample size in a pilot trial [13].

Many EEG based monitoring systems, such as Bispectral Index Scale, Patient State Index, and State Entropy, have been studied at ICU environment for monitoring the depth of sedation. However, according to the FDA, none of them have been directly indicated or validated to measure the level of sedation of ICU patients [66, 67, 68].

For the purpose of monitoring sedation level of ICU patients, GE Healthcare in collaboration with Edinburgh Critical Care Research Group has developed Responsiveness Index. It is a novel measurement parameter quantifying patient responsiveness from frontal EMG (fEMG). The Resposiveness Index is hypothised to be an objective measurement which could help the nurses to titrate sedatives according to patient's needs. However, Responsiveness Index has not yet gained regulatory approval from the FDA, and the pivotal study has not been conducted.

First, this chapter will briefly present the EEG based measures and review the the most notable studies in which they been tested in the ICU for monitoring the level of sedation. Second, the basics of Responsiveness Index are summarized. Third, as this study analyzes the results from a pilot study, common hypothesis tests for trials with small sample sizes are presented.

3.1 Depth of anesthesia monitors

Electroencephalographic (EEG) signals can be used to quantify the patient's level of consciousness. By using mathematical algorithms, the raw EEG signal can be processed to create a single number from 0 (flat EEG) to 100 (fully alert). The current EEG based monitoring systems were primarily developed to monitor depth of anesthesia, not the level of sedation. However, these monitors have been used to assess level of consciousness of ICU patients, [69] as they offer more objective measure of sedation level than clinical scoring systems [70]. These monitors also have the benefit of continuous display of data, which can be used to track unexpected changes in level of consciousness. [3] The following subchapters will present the EEG based monitoring systems which have been tested to measure the sedation level of an ICU patient.

3.1.1 Bispectral Index Scale

Bispectral Index was developed to monitor the level of consciousness in anesthetized patients [71]. In the United States, it has been indicated to be used as an aid in monitoring the effects of certain anesthetic agents. The use with certain anesthetic

agents may be associated with reduction in primary anesthetic use. Additionally, it may help to guide anesthetic administration, thus reduce incidences of awareness with recall during anesthesia and sedation. [66]

BIS has been compared to clinical sedation scales in numerous studies. In a review conducted by LeBlanc and colleagues in 2006, results from 19 of these studies were summarised. They found that the correlation between BIS and clinical assessment scores varied between marginal and good. BIS did not correlate well with clinical assessment of sedation levels in all patients. Additionally, the interindividual differences increased at deeper levels of sedation. The authors recommend further studies to evaluate the impact of the BIS monitoring to patient outcomes. [72]

In 2009, Olson et al. studied the effects of BIS augmented sedation monitoring to patient outcomes in mechanically ventilated neurological patients. The patients were divided into two groups: patients receiving sedation monitoring with BIS and Ramsay Sedation Scale (n = 32), and patients receiving sedation monitoring only with Ramsay Sedation Scale (n = 35). As a result, they found that when BIS was used in conjunction with RSS as opposed to only using RSS, patients received significantly less propofol by volume on average during 12 hour period of monitoring (93.5 ml vs 157.8 ml with $p < 0.015$). Additionally, the patients who received BIS augmented monitoring, had lower infusion rates and woke up significantly quicker. There were no undersedation events during the 12 hour monitoring period within either of the study groups. However, there were no statistical difference noted in Ramsay scores nor in the distribution of the scores between the two groups. [73] The results suggest, that using EEG based sedation monitoring with clinical sedation scales may result in better patient outcomes. However, high electromyographic (EMG) activity has been shown to cause artefacts to the processed EEG values, leading to increased BIS values [74].

3.1.2 Patient State Index

According to the FDA, the Patient State Index (PSI) is an EEG based method intended to measure the state of the brain in the operating room, intensive care unit or research laboratory [67]. As reported by Drover and Ortega, it has been specifically designed to provide an indication of sedation and anesthesia levels during surgery in the ICU. Like BIS, PSI value is derived from EEG signal and it gives a number from the range between 0 (deeply anesthetised) and 100 (fully awake). [75]

In 2003, Schneider and colleagues evaluated whether the PSI indicates the level of sedation as measured by Ramsay score in mechanically ventilated ICU patients. As a result, they found that PSI showed good association with the levels of sedation. [76]

However, not all studies have demonstrated that PSI is a good indicator of sedation level. In 2011, Caputo and the research team evaluated the PSI in a study involving 19 ICU patients whose treatment included surgical therapy with continuous intravenous sedation. In this study, the Ramsay sedation scale and PSI showed poor correlation, partly due to EMG artefacts. In conculsion, the authors suggested that clinical assessment by healthcare practitioner is the best way to monitor patient's sedation level. [77]

3.1.3 State Entropy

State Entropy is an EEG based measure indicated in the US for monitoring the state of the brain and effects of certain anesthetic agents [68]. In 2008, Walsh and colleagues studied whether or not State Entropy is a valid measure of sedation state in critically ill patients by comparing it to RSS. The results showed poor discrimination between different sedation states and in conclusion, the authors stated that fEMG was a major confounder in the study. [78]

3.2 Responsiveness Index

As presented earlier, no device on the market can reliably measure a patient's sedation level. As the devices available were originally designed for measuring the depth of anesthesia, they were adopted or transferred from the anesthesia setting to the ICU. Studies regarding the use of BIS, PSI and State Entropy in the ICU setting suggest that fEMG activity is a major confounder for these algorithms. The reason is that the frequent arousals typical to the ICU patient increase fEMG activity, and the frequency bands of EEG and fEMG overlap. [79]

Responsiveness Index uses a different method than EEG based measures to assess the level of sedation. It is a parameter which quantifies the patient's responsiveness from fEMG. It is hypothized that fEMG is useful for measuring the sedation level of the patient, as it measures the patient's responsiveness caused by the stimulation during treatment.

3.2.1 Responsiveness and frontal electromyography

Human body creates various responses to changing emotional states and the face plays a substantial role in expressing these states [80]. Even though consciousness is a cortical phenomenon, the facial muscle activity is controlled by the brainstem [81] (see figure 4).

The electrical activity of muscle contractions can be measured with electromyography (EMG) [82]. Frontal electromyography (fEMG) can be used to detect patient's responsiveness to ambient and internal stimuli by measuring the electrical activity of frontal facial muscles with three electrodes (see figure 5). Edmonds and colleagues have demonstrated that painful and/or stressful stimuli increases fEMG amplitude in conscious and lightly anesthetized patients, [83] and sudden fEMG activity has been associated with increased patient responsiveness [84].

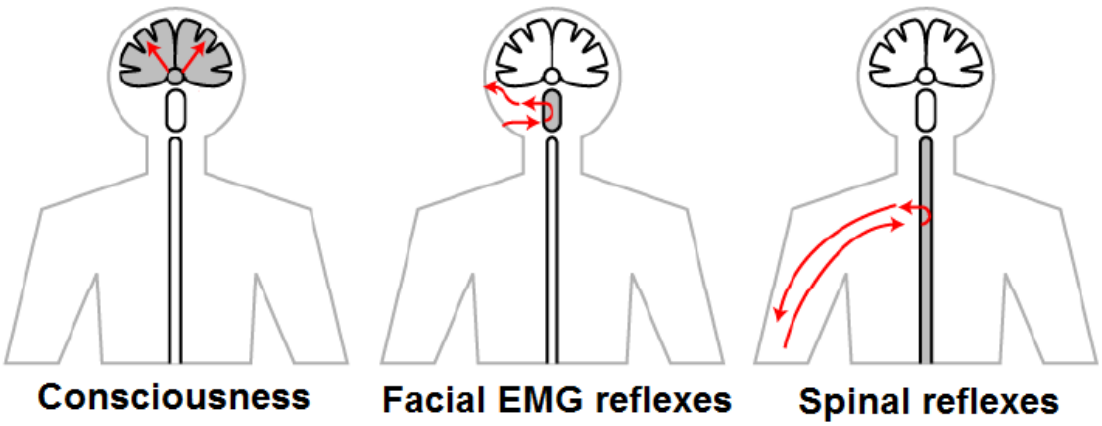


Figure 4: Wakefulness and facial muscle responses are controlled by the brainstem, whereas consciousness requires cortical activity, and reflex movements are controlled at spinal level.

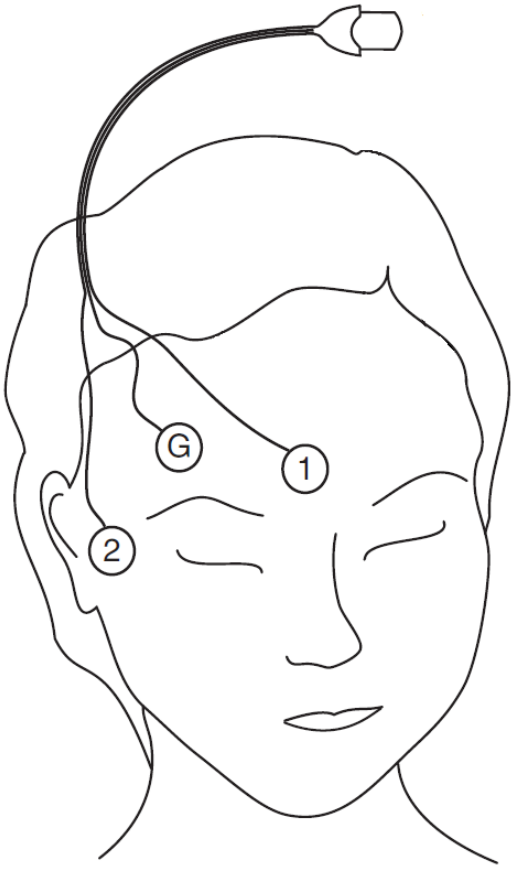


Figure 5: Electrode setup. Numbers 1 and 2 denote the active electrodes; G represents the ground electrode. Modified from publication by Lapinlampi et al. [85].

The fEMG signal is affected by the following factors in ICU setting: 1) the intensity and frequency of stimuli caused by patient care and the underlying condition; 2) the sedative and analgesic drug received by the patient; and 3) the effect of the illness and drug treatment on brain and muscle functions, specifically encephalophathy, delirium and use of neuromuscular blocking agents (NMBAs), which may alter the fEMG responses to stimuli. [85]

3.2.2 Algorithm

Responsiveness Index algorithm quantifies the fEMG responses in relation to ambient (light, noise, care procedures) and internal (pain, agitation) stimuli by giving out a single number from 0 (unresponsive) to 100 (fully responsive) describing the strength of the response. As the ambient and internal stimuli varies over time in the ICU patients, the RI is hypothized to be a dynamic variable representing the balance between stimulation and sedative drug effect [79]. No active stimulation is required, because the patients are constantly stimulated in the ICU environment by light, noise, and care procedures.

The RI algorithm development was performed in three stages. First, the fEMG activity related to standard vocal stimulus in ICU patients was characterized. The vocal stimulus used was a command of "open your eyes" with 80-dB intensity. The response was considered positive if the patient opened his or her eyes within 10 seconds after the stimulus. The epochs containing fEMG waveform starting 2 minutes before and ending 2 minutes after the stimulus were extracted. The epochs were divided to 0.5-s windows and estimate of the fEMG power was calculated. The fEMG powers were analyzed separately for patients who had positive response and negative response. Median and the interquartile ranges were calculated from signal powers to describe the form of typical fEMG response to vocal stimulus. After this, the algorithm was developed in iterative process by using the characterized data. Finally, the algorithm was evaluated against modified Ramsay Sedation Scale which used standardized vocal, loud vocal (similar to standardized, but louder), and tetanic stimuli. [85]

The algorithm calculates root mean square power in 5-second epochs from fEMG using 50–150-Hz frequency band. Before the power calculation, fEMG signal is filtered with high-pass finite impulse response filter and 10-Hz comb filter, which attenuates the mains frequencies and 10-Hz multiples. The filtering removes artefacts caused by movement and blinks, which operate on low frequencies. Then, the power value time series is processed with a filter that extracts the patient response related steep rises. The RI value at time t can be expressed with the following equation:

$$RI(t) = S\left(\sum_{n=t-N}^{t} g(n) log\left(\frac{P_{\mathrm{F}}(n)}{1\mu\mathrm{V}} + 1 \right) \right), \tag{1}$$

in which the filtered power value at time point t is marked with P_{F}. The scaling function is denoted with S, which scales the RI value to range between 0 to 100. The weighting function g operates as a low-pass filter and gives more weight to recent fEMG changes. The 1μ sets the minimum value of logarithm to zero. N is the number of P_{F} values in the summation and it has been set to 720 samples, which

corresponds to 1 hour. [85] The reason for setting $N = 720$ samples was, that in order to measure the responsiveness level of the patient in a reliable manner, a longer measurement period was needed.

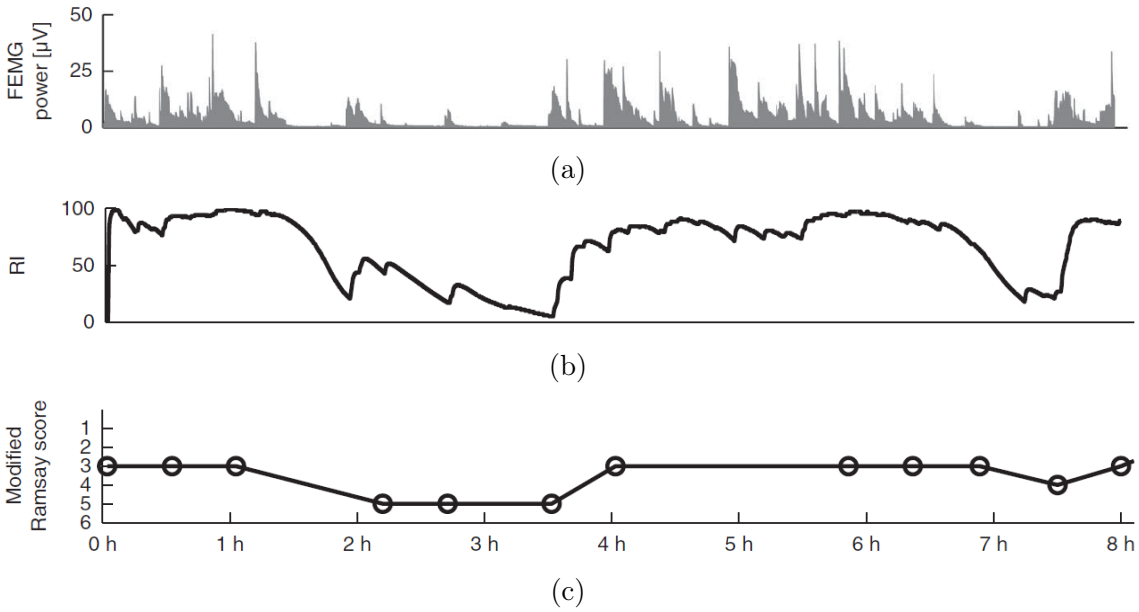


Figure 6: (a) The fEMG power, (b) the RI value in figure, and (c) the corresponding modified Ramsay score. Modified from publication by Lapinlampi et al. [85].

3.2.3 Intended use

Responsiveness Index is not intended to replace the role of clinicians in making patient specific sedative administration decisions in the ICU. It is also not designed to replace the current sedation practices. The clinical benefit of RI may be that it provides additional information to the clinician and supports the decision making process of administration of sedatives when used as an adjunct to other parameteres.

RI measurement has some limitations. The RI value is subject to individual patient differences, and therefore, the RI target level for optimal sedation should be adjusted individually for each patient. In addition, RI measurement does not work with patients that are receiving neuromuscular blocking agents (NMBAs), since the facial muscle reactions are blocked. In these cases, the RI value will remain low and might give the user an inaccurate view of patient's sedation status. As the drug effect is terminated, the measurement starts working reliably again.

3.3 Hypothesis testing

Pilot trials are not used as conclusive evidence to support the safety and efficacy claims of the device. The objective of a pilot trial is to study the feasibility of an approach that is intended to be used in a pivotal study. [12, 13]

As the sample sizes in pilot trials are small, they are not ideal for hypothesis testing [12]. However, pilot studies can be used to look for trends, but the results should be interpreted with caution. The statistical testing for continuous variables in a pilot trial can be performed with non-parametric methods, such as Wilcoxon rank-sum test. For categorial data, Fisher's exact test can be used to explore the association between the categories when sample sizes are small. Kaplan–Meier curve with Logrank test is a viable option for time-to-event analysis when censoring is required. These tests are briefly presented in the following subchapters.

3.3.1 Wilcoxon rank-sum test

Wilcoxon rank-sum test can be used to perform non-parametric analysis to test the null hypothesis of variables which are not normally distributed. The null hypothesis is that the samples come from the same population. [86]

First, Wilcoxon rank-sum test ranks all the observations as if they were from a single sample. The smallest value has the rank value 1, and 2nd smallest rank value 2, and so on. Then, the sum of the ranks is calculated for both groups. By taking account the number of samples in both groups and the sum of ranks of the group where n is smaller, the p-value can be defined from Wilcoxon rank-sum distribution table. [86]

3.3.2 Fisher's exact test

Fisher's exact test is used for hypothesis testing of contingency tables (see table 8) when sample sizes are expected to be small. It explores the association of categorial data of paired observations by giving out a p-value. The null hypothesis tested is that there is no association between the groups and the categories.

Table 8: Example of a 2 × 2 contingency table.

	Trial	**Control**	Row total
Yes	a	b	$a + b$
No	c	d	$c + d$
Column total	$a + c$	$b + d$	$a + b + c + d$

The propability of obtaining a set of values with cell frenquencies a,b,c, and d when null hypothesis is true is given by

$$p = \frac{(a+b)!(a+c)!(b+d)!(c+d)!}{N!a!b!c!d!} \tag{2}$$

where N is the sum of cell frequencies a, b, c, and d.

3.3.3 Kaplan–Meier curve and Logrank test

Kaplan–Meier curve illustrates the propability of occurence of defined event as a function of time. It can be used to compare the time-to-event of patients between

the groups. The curve is plotted as a step function, in which each step corresponds to one patient's time-to-event. The events can be censored to indicate that the time of event occurence cannot be accurately determined. Censoring is done, when the required data are not available or the study period ends before the subject went through the defined event. [86]

The comparison of statistical difference of the Kaplan–Meier curves can be performed by using Logrank test. It is a non-parametric method for testing whether or not the groups are samples from same population with respect to patients' time-to-event. The method ranks all event occurence times (excluding censored), and produces an observed (O) and expected (E) number of events for both groups. Using these variables, the chi-square (X^2) can be calculated, which is a measure of the degree of deviation between the observed and expected result:

$$X^2 = \frac{(O_\mathrm{a} - E_\mathrm{a})^2}{E_\mathrm{a}} + \frac{(O_\mathrm{b} - E_\mathrm{b})^2}{E_\mathrm{b}}, \tag{3}$$

in which a signifies values of group a, and b signifies values of group b. By comparing the result of X^2 to chi-squared distribution χ^2 with one degree of freedom, the p-value can be defined. [86]

4 Materials and methods

A randomized controlled pilot trial was performed to assess the effectiveness and safety of continuous RI monitoring during early ICU care as nurse decision-support tool. The study took place in an 18-bed Scottish general adult ICU admitting approximately 650 ventilated medical, surgical, and trauma patients annually (excluding routine cardiac surgery and neuro-intensive care). Patients requiring mechanical ventilation and sedation were randomized via sequential sealed envelopes following ICU admission. The trial group patients received RI-augmented sedation monitoring and control group patients were treated according to current practice. In trial group, the RI monitoring was visible and the caregiving staff were asked to adjust sedation to maintain a state where RI ≥ 20 and patient responds to verbal stimulation. In control group, the RI monitor was connected to the patients but the data were concealed from clinical staff.

For the RI monitor to demonstrate proof of concept and acceptability, the patients who received RI monitoring were expected to spend more time with higher RI values. To explore this hypothesis, the trial group patients were compared to control group patients in terms of proportion of RI < 20. In addition, the groups were compared in terms of incidences of deep sedation (proportion of time with RASS ≤ -4), the time to reach RI ≥ 20 and RASS > -4, and patient outcomes. The patient outcomes were defined as the number pre-defined adverse events; the time to reach first extubation; the administered dose of sedative and analgesic drugs; the use of sedation holds; ICU-LOS; and number of deaths.

In the secondary analysis, two patient subgroups were analyzed. The first subgroup consisted of patients with RI < 20 at the start of monitoring and the second subgroup consisted patients who were deeply sedated at the start of monitoring. The hypothesis in the secondary analysis was that the effects of RI monitoring would be more visible in these subgroups, as the nurses were only asked to alter the sedation when RI was < 20 and the patient did not respond to verbal stimulus.

4.1 Patient enrollment and randomization

A total number of 90 ICU patients were enrolled into the study. The patient data set was gathered in Royal Infirmary of Edinburgh ICU between 2009 and 2010. Patients were screened by the clinical nursing staff with following inclusion criteria:

1. patients mechanically ventilated via endotracheal tube; and
2. patients receiving intravenous sedation with a hypnotic agent (midazolam or other benzodiazepine) or propofol by continuous infusion.

Only mechanically ventilated patients were considered eligible, because excessive sedation is common among these patients [79]. The exlusion criteria for this study were

1. primary intracerebral disorder;
2. head injury (causing reduced conscious level prior to intubation);

3. patient who was already conscious at the time of enrolment defined as RASS ≥ -1;
4. less than 16 years of age;
5. patient not expected to survive the next 24 hours;
6. patient was receiving long term ventilation prior to ICU admission;
7. patient had a long term tracheostomy prior to ICU admission;
8. patient transferred, sedated and mechanically ventilated from another ICU (unless recruitment possible within 24 hours of first ICU admission);
9. patients receiving continuous neuromuscular blocking agent at the time of screening;
10. Previous enrollment in the trial;
11. status epilepticus;
12. confirmed meningitis or encephalitis; and
13. a known chronic neurological disease interfering with normal neuromuscular functions.

The patients were divided into trial group (n = 44) and control group (n = 46) at time of ICU admission. Patients in trial group received care based on the hospitals current practice plus responsiveness monitoring, while patients in control group only received care based on current practice. The current practice was intended to be consistent with recommended best practice, including daily sedation holds and use of a clinical sedation scale linked to decision-making strategy.

A major focus of this study was in the first 48 hours of intensive care, so the patients were entered to the trial from the time of ICU admission and the consent from relatives was obtained retrospectively. The patients were randomized to trial and control group by using sealed envelopes immediately after the ICU admission.

4.2 RI monitoring in trial group

Responsiveness Index monitor was attached to the patient and data was presented to the clinical staff continuously. All participating nursing staff received pre-trial training in the study protocol and to the use of RI monitor.

The monitor presented a continuous trend over time that was color coded using a traffic light system. A RI number was also recorded representing the most recent value in a separate window. The monitor gave instructions based on the current RI value. An example of a monitoring window is shown in table 9. Prompts accompanied the color of RI presented on the screen, to encourage sedation reduction for patients with a low RI values.

The pre-trial training included introduction to the concept that low RI values are expected for deeply sedated patients, as well as some cases with coma unrelated to sedation, and during natural sleep. Nurses were also instructed not to use the monitor data when patients received neuromuscular blocking agents. Nurses were asked to alter sedation using clinical judgement to transition patients out of the red RI range and adjust sedation to achieve values in the amber/green range.

Table 9: RI monitoring screen.

Monitor color	Monitor instruction on screen	Corresponding RI value
Green	Ensure adequate analgesia Continue current sedation unless patient agitated	> 40
Amber	Moderate risk of oversedation **Text below box:** Ensure adequate analgesia	20–40
Red	High risk of oversedation **Text below box:** Ensure adequate analgesia if responsive to stimuli, eg suctioning/physiotherapy Reduce sedation dose if no eye-opening to physical stimuli	< 20

4.3 RI monitoring in control group

All patients in control group were attached to responsiveness monitor, but the data were blinded from clinical staff. RI data could only be accessed by code, which was not available to the care providers.

4.4 Conditions to disconnect or reattach responsiveness monitor

Responsiveness Index monitoring was used continuously until one of a pre-defined set of criteria to discontinue was met. The monitoring was discontinued for the following reasons:

1. 48 hours were elapsed from ICU admission or from first intubation;
2. the patient was extubated;
3. the patient died;
4. the patient was transferred to another ICU;
5. relative consent was declined; or
6. decision was made to withdraw from treatment.

Clinical staff was able to discontinue RI monitoring during care procedures or other reasons according to clinical judgement. In these cases, monitoring was reattached as soon as feasible.

The aim was to use RI monitoring whenever the patient was intubated and intravenous sedation was administered. The monitoring was continued during temporary sedation stops that were part of the standard care procedure.

4.5 Recording RASS scores in trial and control group

The care giving nurse performed hourly sedation assessment by using RASS scoring. The results were entered to the patient's clinical research file. If the nurse was not

able to perform the scoring or the patient was sleeping, then RASS scoring was not performed in that particular hour. No trial specific instructions regarding sedation titration was linked to RASS scoring and nurses were able to modify dosing based on forms of routine and part of usual practice. RASS scoring was continued throughout the monitoring period unless the patient died or a decision to withdraw treatment was made or the patient was discharged from the ICU.

4.6 Daily data collection

Besides continuous recording of RASS scores and RI values, the following data were collected on a daily basis:

1. Total sedative dose:
    - propofol
    - midazolam
    - other
2. Total opiate dose:
    - morphine
    - alfentanyl
    - other
3. Use of formal sedation holds
4. Extubation times
5. Occurence of following pre-defined sedation related adverse events:
    - unplanned extubation (self-extubation)
    - unplanned removal of vascular catheter
    - unplanned removal of nasogastric/enteral tube
    - episode of myocardial ischaemia
    - myocardial infarction
    - episode of agitation requiring bolus treatment (rescue medication)

4.7 Preprocessing RI values

Periods where Neuromuscular blocking agents (NMBAs) were administered plus 30 minutes after were removed from RI data if the monitoring was on. NMBAs were administered to 13 patients. One patient received NMBAs in two occasions. In 10 patients, RI monitoring was on when the NMBAs were administered (n = 6 in the trial group and n = 4 in the control group).

Table 10: Occurence of NMBA administration during the intervention period and removal of RI data.

Trial group					
Patient ID	Date	Start time	Stop Time	Monitor status	Removed period
002	2009-12-10	15:45	15:45	Monitoring ended	Nothing removed
003	2010-01-08	13:00	13:30	Monitor attached	13:00 – 14:00
020	2010-03-05	03:05	03:08	Monitor attached	03:05 – 03:38
020	2010-03-05	05:50	05:53	Monitor attached	05:50 – 06:23
027	2010-03-10	17:20	17:20	Monitor attached	17:20 – 17:50
033	2010-03-20	14:00	16:30	Monitor attached	14:00 – 17:00
045	2010-04-09	16:15	Not reported†	Monitoring not started	Nothing removed
072	2010-06-13	23:05	23:10	Monitor active	23:05 – 23:40
081	2010-06-30	13:30	13:30	Monitor active	13:30 – 14:00
Control group					
021	2010-03-05	22:00	22:05	Monitoring not started	Nothing removed
024	2010-03-09	17:15	17:15	Monitor attached	17:15 – 17:45
040	2010-04-04	21:00	21:30	Monitor attached	21:00 – 22:00
043	2010-04-06	15:45	15:50	Monitor attached	15:45 – 16:20
044	2010-04-09	15:40	15:50	Monitor attached	15:40 – 16:20

† Stop time of NMBA administration was not added to patient's clinical file, but it was noted that NMBA administration stopped before the monitoring started.

4.8 Preprocessing RASS scores

RASS scores were preprocessed with the following algorithm:

1. If RASS scoring was not performed during a particular hour, the next RASS value was included in the analysis.

2. In a situation where the next RASS score did not exist within one hour, the RASS score from previous hour was taken.

3. If both, the next and the previous RASS values were missing, then the RASS value of that specific time point was considered as missing.

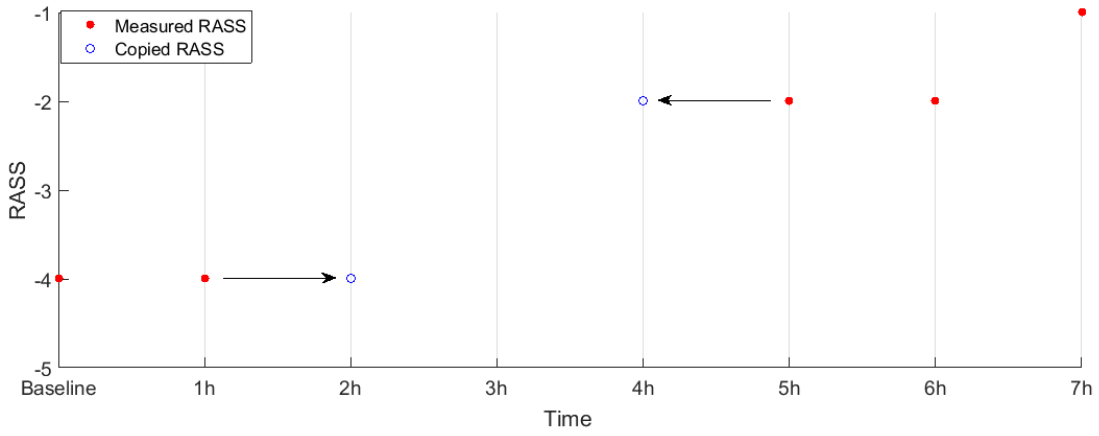


Figure 7: Illustration of RASS score preprocessing. At time points 2h and 4h, RASS scoring was not performed, and the neighbor RASS score was copied. At time points 2h, 3h, and 4h, RASS scoring was not performed, thus time point 3h was considered as missing.

4.9 Administration of drugs during study

The clinical team made decision of choise and administration of sedatives and analgesic drugs. The standard sedative drug was propofol, and analgesic was alfentanil, at the time of the study. Clinicians used alternative agents, principally midazolam and morphine, according to clinical discretion. All decisions about other aspects of usual care were made by the clinical team. The only trial-specific interventions related to changes in the dose of sedatives was made by the bedside nursing staff.

4.10 Analysis

The effects of Responsiveness monitoring during the intervention period and after 7 days were explored by using several methods. The intervention period was defined as the time from start of Responsiveness monitoring until 48 hours had elapsed from ICU admission.

Primary analysis was performed for all patients. The secondary analysis was performed for patients with baseline RI < 20 and for patients with baseline RASS ≤ -4. The RI value at baseline was defined as the first valid RI value (RI value after 30 minutes from start of monitoring). RASS score at baseline was defined as the first RASS score after the start of RI monitoring.

All the analyses were performed using MATLAB (version R2014b) with Statistics Toolbox and R (version 3.2.1).

4.10.1 Baseline differences

In order to test baseline differences between the groups, the following data were recorded:

1. age;
2. gender;
3. time from intubation to start of monitoring;
4. RI value at baseline;
5. RASS score at baseline; and
6. total intravenous sedation and analgesic drug dose prior to monitoring.

The patient groups were also compared at baseline in terms of severity-of-illness scores. The following scoring systems were used: Acute Physiological and Chronic Health Evaluation (APACHE II score), Charlson Comorbidity Index, and Sequential Organ Failure Assessment (SOFA) score. In this study, SOFA scoring for central nervous system was omitted because in sedated patients it is unreliable. Therefore, the maximum SOFA score was 20.

4.10.2 RI values and RASS scores

Since RI value measures the patient responsiveness to ambient and internal stimuli, RI values were expected to be lower during night when patients are asleep and less care procedures are performed by the caregiving staff. Additionally, the RASS scores were expected to be lower during the night as the nurses may administer more sedatives to help the patients to sleep. In this study, night time was defined as the time between 22:00-08:00 and day time as the time between 08:00-22:00. First, RI values and RASS scores recorded at night time were compared to the RI values and RASS scores recorded at day time. Second, the groups were analyzed in terms of whether or not they had similar proportions of day and night RI values and RASS scores. As RI uses 720 samples (corresponding to 1 hour) to calculate the RI value, therefore the RI values are not independent of each other. Hence, in this comparison, a RI sample from each hour from each patient was used. This analysis was done in order see if the comparison of groups was affected by the possible day and night variation.

Descriptive analyses for the time evolution of RI values and RASS scores were performed to compare the differences between the two groups in defined time points during the intervention period. The motivation for the time evolution analyses was to find the time period when the RI monitoring was most effective. Two factors were identfied which can distort the time evolution analyses of RI values and RASS scores:

1. RI values and RASS scores were expected to vary between day and night.

2. RI values and RASS scores were expected to be lower during early phase of monitoring. The reason is that RI monitor was attached as soon as possible after the patient was intubated. Patients are expected to require more sedatives in the early phase of the mechanical ventilation when endotracheal tube has recently been installed.

Therefore, the time evolution analyses of RI values and RASS scores were performed using two different methods. In method one (time of the day RI and RASS evolution), the day and night variation of RI values was taken into account when

comparing the values between the groups. The analysis was done as follows for the RI values:

1. The intervention period was divided into four-hour time intervals ranging from 00:00 to 24:00.

2. RI samples were selected from each patient every four-hours after recruitment until the end of the intervention period.

3. The number of patients in each RI category was calculated in each four-hour time interval. The categories were red (RI $<$ 20); amber (20 $\leq$ RI $\leq$ 40); green (40 $<$ RI $\leq$ 100); and RI monitor off and intubated. Additionally, the number of patients in whom RI monitoring discontinued due to extubation or death were included in the analysis. This means that there were six different categories which were RI $<$ 20; 20 $\leq$ RI $\leq$ 40; 40 $<$ RI $\leq$ 100; Extubated; Died; and RI monitor off and intubated.

4. A stacked histogram was created to compare number of patients in each category in given times of the day. As the intervention period lasted up for 48 hours, and first RI value was taken between 00:00 and 24:00, the histogram covers RI values for a period of 72 hours.

A histogram was also created for the RASS scores taken during the intervention period by using the same method, but without adding the number of patients extubated. The categories for RASS were RASS $\leq$ –4; RASS = –3; 0 $\geq$ RASS $\geq$ –2; and RASS $\geq$ 1.

In method two, the time evolution of the RI values and RASS scores during the intervention period were analyzed as a function of time elapsed from start of monitoring. A sample from every four-hours from each patient was used. For RI values, a box plot and stacked histogram were created to allow the comparison of the values between the groups in each time point. The time evolution of RASS scores were illustrated with a histogram showing the number of patients with each RASS score every four-hours during the intervention period. Additionally, a histogram showing the proportion of RI and RASS samples taken during night in each time point was created.

The groups were also compared in terms of time to reach first RI $\geq$ 20, and time to reach first RASS $>$ –4 by using the Kaplan–Meier curve.

Finally, the differences between the groups at the end of the intervention period were analyzed with respect to proportion of time spent on each RI category and extubated, and proportions of RASS scores.

4.10.3 Patient outcomes

The groups were also compared in terms of patient outcomes. The compared patient outcome variables after the intervention period were

1. total sedative and analgesic drug doses;

2. number of patients who went through sedation hold;
3. number of patients extubated;
4. number of patients who died; and
5. number of pre-defined adverse events.

The compared patient outcome variables after 7-days following randomization were

1. time to first extubation;
2. number of patients extubated;
3. number of patients who died;
4. number of patients discharged from the ICU;
5. number of patients still in the ICU;
6. number of patients still in the ICU receiving mechanical ventilation; and
7. number of patients transferred to another ICU.

4.10.4 Secondary analysis

In the secondary analysis, only the patients who had baseline RI < 20, and the patients who had baseline RASS $>$ -4 were analyzed. The comparison of trial and control group was performed in terms of

1. proportion of RI < 20 and RASS ≥ -4 after the intervention period;
2. RI values and RASS scores at baseline; and
3. total sedative and analgesic drug doses during the intervention period.

4.10.5 Hypothesis testing

Due to small sample size, the continuous variables were not expected to be normally distributed. The hypothesis testing for the non-parametric variables was performed with Wilcoxon rank-sum test. Fisher's exact test was used for all categorial data. Survival analysis using Kaplan-Meier curve and Logrank test were performed for time-to-event variables. A $p < 0.05$ was considered statistically significant.

5 Results

5.1 Baseline characteristics

A total number of 74 patients (n = 36 in the trial group and n = 38 in the control group) out of 90 were included in the analysis. 16 patients were excluded from the study due to following reasons: no consent (n = 11), died before consent (n = 1), missing envelope (n = 2), or the monitoring was not started (n = 2). Patients' charasteristic data are listed in Table 11. The patient groups were similar at baseline with respect to all measured variables.

Table 11: Patients' characteristics at baseline.

	Group		Hypothesis test
Baseline variable	Trial	Control	p-value
Number of patients	36	38	NA
Age in yrs, median (1st, 3rd quartile; min-max)	60 (44, 69; 25 – 85)	59 (43, 72; 27 – 80)	0.725^{W}
Sex, male/female	21/15	26/12	0.469^{F}
APACHE II score, median (1st, 3rd quartile; min-max)	20 (11, 24; 0 - 31)	23 (17, 26; 0 - 38)	0.083^{W}
Charlson Comorbidity Index, median (1st, 3rd quartile; min-max)	1 (0, 2; 0 - 5)	1 (0, 2; 0 - 7)	0.835^{W}
SOFA score from the first 24 hours, median (1st, 3rd quartile; min-max)	7 (5, 11; 2 - 16) *n=35	8 (5, 9; 1 - 16) *n=37	0.847^{W}
RASS score at start of intervention period, median (1st, 3rd quartile; min-max)	-3 (-4, -2; min= -5, max=0)	-4 (-4, -3; min= -5, max=2)	0.695^{W}
Number of patients who had valid RASS score at study entry, n (%)	27 (91.7%)	31 (92,1%)	0.578^{F}
RI at start of intervention period, median (1st, 3rd quartile; min-max)	16 (0, 55; 0 – 100)	18 (0, 42; 0 – 100)	0.589^{W}
Red, n (%)	19 (54.3%)	20 (54.1%)	1.000^{F}
Amber, n (%)	5 (14.3%)	8 (21.6%)	0.545^{F}
Green, n (%)	11 (31.4%)	9 (24.3%)	0.604^{F}
No value, n (%)	1 (2.8%)	1 (2.6%)	1.000^{F}
Time (in hrs) from intubation to start of monitoring, median (1st, 3rd quartile; min-max)	4.0 (2.1, 6.2; 1.0 - 12.0)	4.9 (3.5, 7.7; 0.7 - 11.8)	0.193^{W}
Total propofol dose (in mg) prior to monitoring, median (1st, 3rd quartile; min-max)†	205 (140, 720; 0 - 7200)	430 (170, 780; 0 - 2360)	0.327^{W}
Total alfentanil dose (in mg) prior to monitoring, median (1st, 3rd quartile; min-max)‡	1.75 (0.75, 5.25; 0.00 - 14.00)	3.00 (1.00, 7.50; 0.00 - 17.50)	0.190^{W}

^{W} Wilcoxon rank-sum test p-value

^{F} Fisher's exact test p-value

* SOFA scoring was not performed for all the patients during the first 24 hours of ICU care.

† 1 mg of midazolam was considered equivalent to 10 mg of propofol. Two patients received midazolam (n = 0 in trial group and n = 2 in control group).

‡ 1 mg of morphine was considered equivalent to 100 μg alfentanil. One patient in control group and none of the patients in trial group received morphine.

5.2 RI values

5.2.1 Day and night analysis

Comparison of day and night RI values (Figure 8) shows tha thet RI values were lower during night time when less care procedures were performed and patients were sleeping. The median RI value during day time was 29 (n = 1246) and during night time 19 (n = 889). The difference was statistically significant (Wilcoxon rank-sum test $p < 0.001$). According to Table 12, both groups had similar proportions of day and night RI values.

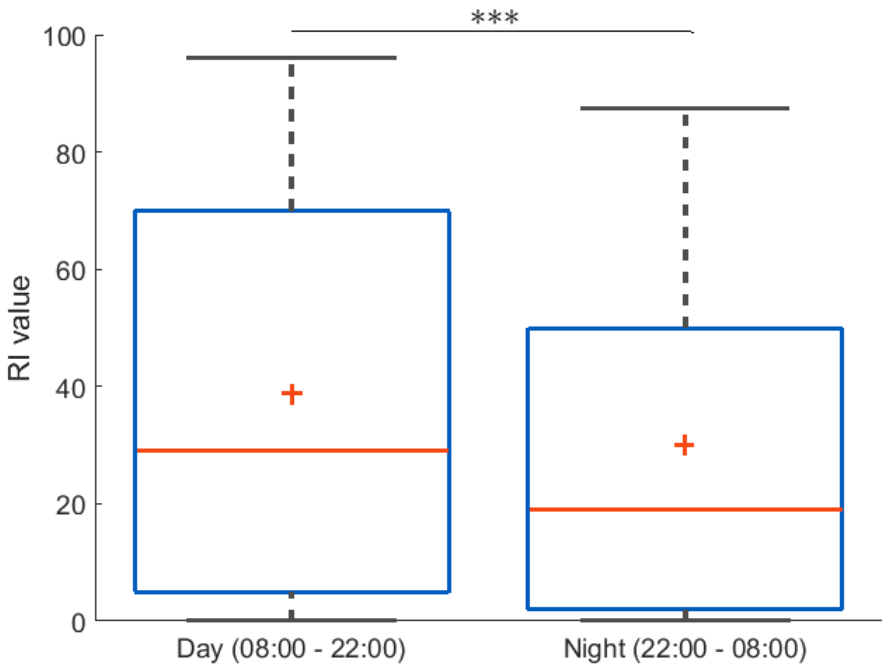


Figure 8: Difference in day and night RI values. The values were independent, as one RI sample from each hour from each patient was used. Cross-line within box indicates median and + value is mean, while the upper box limit is 3rd quartile and lower box limit is 1st quartile. Whiskers represent the range between 9 – 91%. A p-value < 0.001 is denoted with ***.

Table 12: Proportions of day time and night time RI values for both groups.

	Day (08:00-22:00)	**Night (22:00-08:00)**
Trial group	57.1%	42.9%
Control group	57.9%	42.1%

5.2.2 Time-of-day RI evolution

Figure 9 compares of RI values between the groups at different times of the day. The figure also displays if the patient was extubated or the RI monitor was off while the patient was intubated. None of the patients died during the intervention period. The first day shows the patient enrollment to the study. The valid data for comparison start from the beginning of the second day, when all patients had been recruited. The number of patients is declining during the third day, as 48 hours had elapsed from ICU admission.

According to the same figure, trial group had less patients with red RI values (RI $<$ 20) compared to control group in all time intervals during the second day. The proportions of patients in both groups with RI $<$ 20 in each time interval during the second day are listed in Table 13. Additionally, the figure indicates that the patients in the trial group were extubated faster during the intervention period.

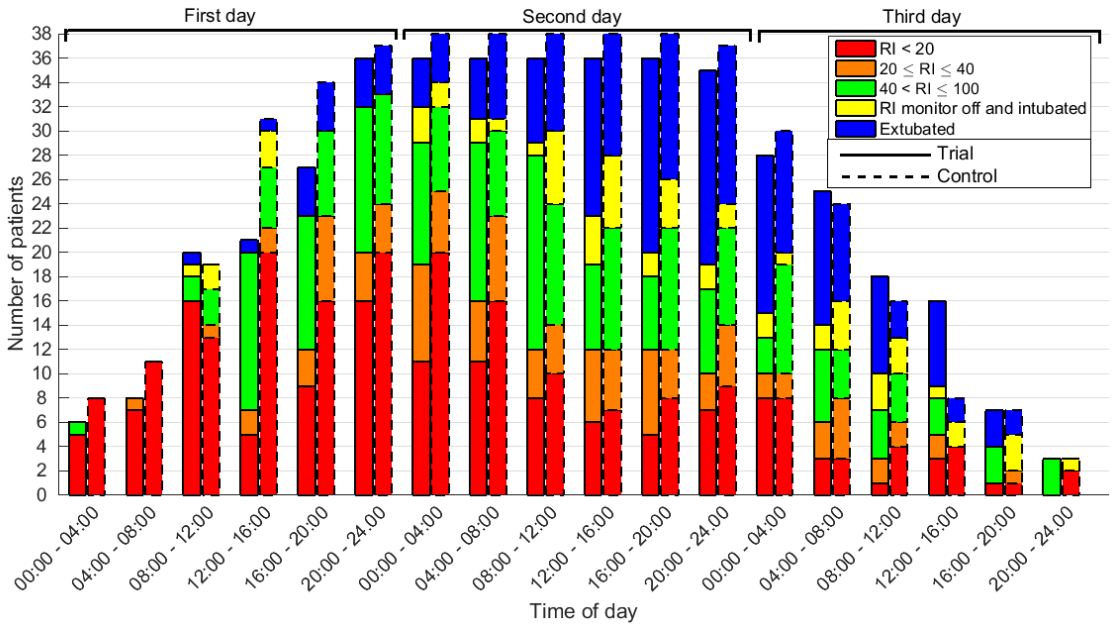


Figure 9: Time-of-day RI evolution. A RI sample was selected from each patient every four hours after the start of monitoring. A maximum of 48 hours of data were used for all patients. None of the patients died during the intervention period. Left-hand side values belong to the trial group (solid bar lines) and right-hand side values to the control group (bars outlined with '- -').

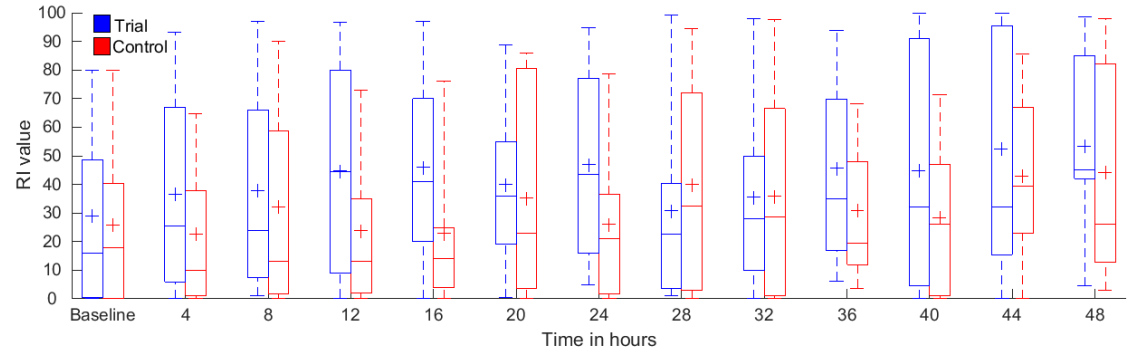
Table 13: Proportions of patients with RI < 20 during the second day in each time interval.

Time of day	**Trial group (%)**	**Control group (%)**
00:00-04:00	31.4	52.6
04:00-08:00	30.5	42.1
08:00-12:00	22.2	26.3
12:00-16:00	16.6	18.4
16:00-20:00	13.9	21.1
20:00-24:00	20.0	24.3

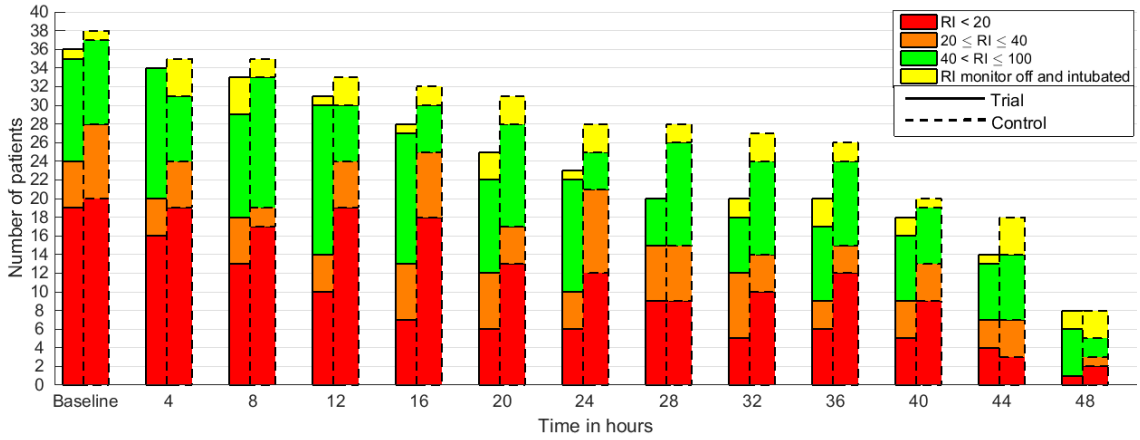
5.2.3 RI evolution as a function of time elapsed from start of monitoring

In Figures 10a and 10b, the evolution of RI values are illustrated as a function of time elapsed from start of monitoring in four-hour time intervals. Figure 10c shows the proportions of RI values taken at night in each time point.

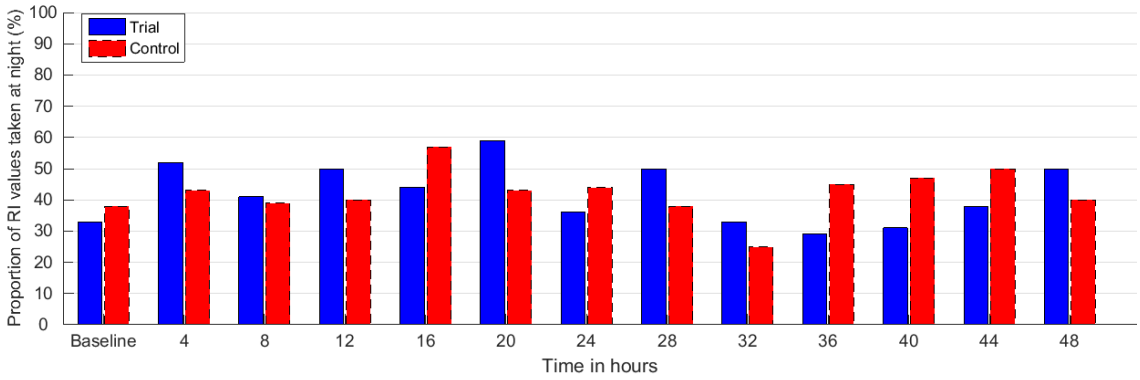
Both Figures (figures 10a and 10b) show that the patients in the trial group had higher RI values compared to the control group. The most notable difference between the groups can be seen from baseline up until 24 hours has elapsed from start of RI monitoring. Figure 10c does not indicate any major differences in the proportions of RI values taken at night between the groups. Therefore, the higher RI values in the trial group cannot be explained with the day and night variation of RI values.



(a)



(b)



(c)

Figure 10: (a) RI time evolution box plot from start of monitoring to 48 hours. Cross-line within box indicates median and + value is mean while the upper box limit is 3rd quartile and lower box limit is 1st quartile. Whiskers represent the range between 9-91%. (b) RI time evolution histogram from start of monitoring to 48 hours. Number of patients is declining due to following reasons: decision was made to withdraw patient from treatment, patient was extubated, or 48 hours had elapsed since ICU admission or intubation. Left-hand side values belong to the trial group (solid bar lines) and right-hand side values to the control group (bars outlined with '--'). (c) Proportion of RI values taken at night.

5.2.4 Time to reach first RI $\geq$ 20

Figure 11 (Kaplan–Meier curve) shows the time from start of monitoring to first RI $\geq$ 20. According to the figure, patients in trial group reached first RI $\geq$ 20 slightly faster than control group patients. However, the difference was not statistically significant (Logrank test $p = 0.428$).

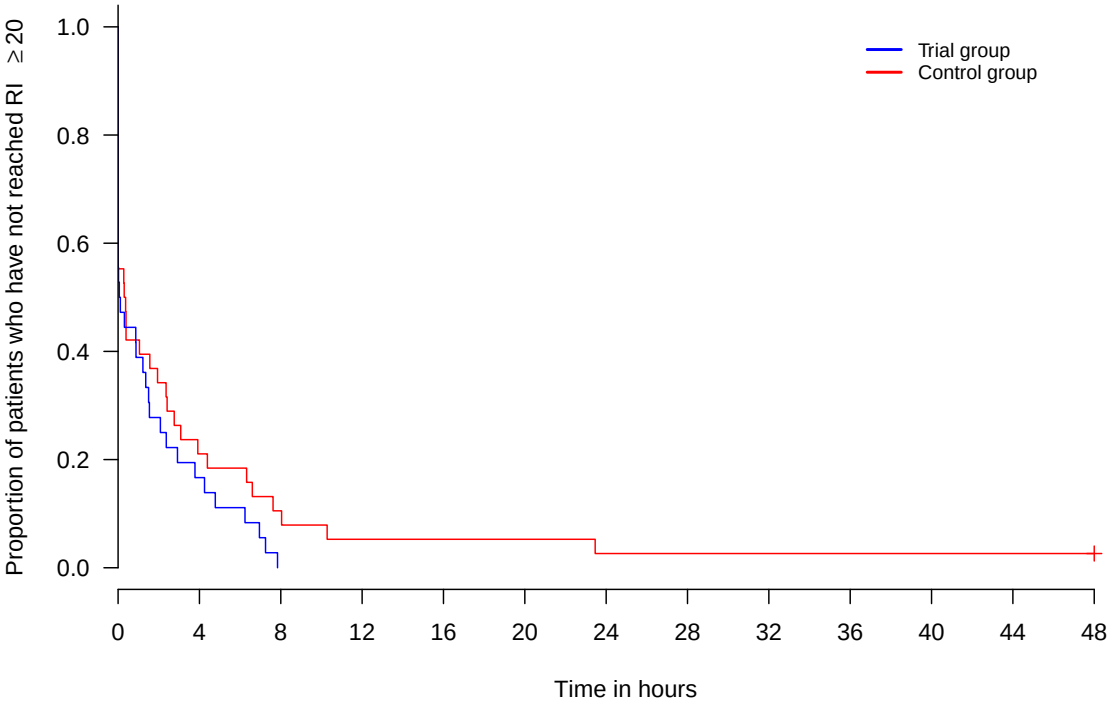


Figure 11: Kaplan-Meier curve of time to reach first RI $\geq$20. Censored events are marked with cross. One patient in control group never reached RI $\geq$ 20, and therefore was censored. A total of 35 patients (n=17 in trial group and n=18 in control group) were censored, because their baseline RI value was $\geq$ 20.

5.2.5 Proportions of RI values after the intervention period

The proportions of time the patients spent on each RI category or extubated after the intervention period are shown in Figure 12. The primary category of interest was the proportion of time spent with RI < 20, and the other categories were plotted for the sake of curiosity.

The figure shows that patients in trial group spent less time with RI < 20 (median of 15.6% vs. 33.4%). However, the result was not statistically significant (Wilcoxon rank-sum test $p = 0.077$), but can be considered as a clear trend towards significance.

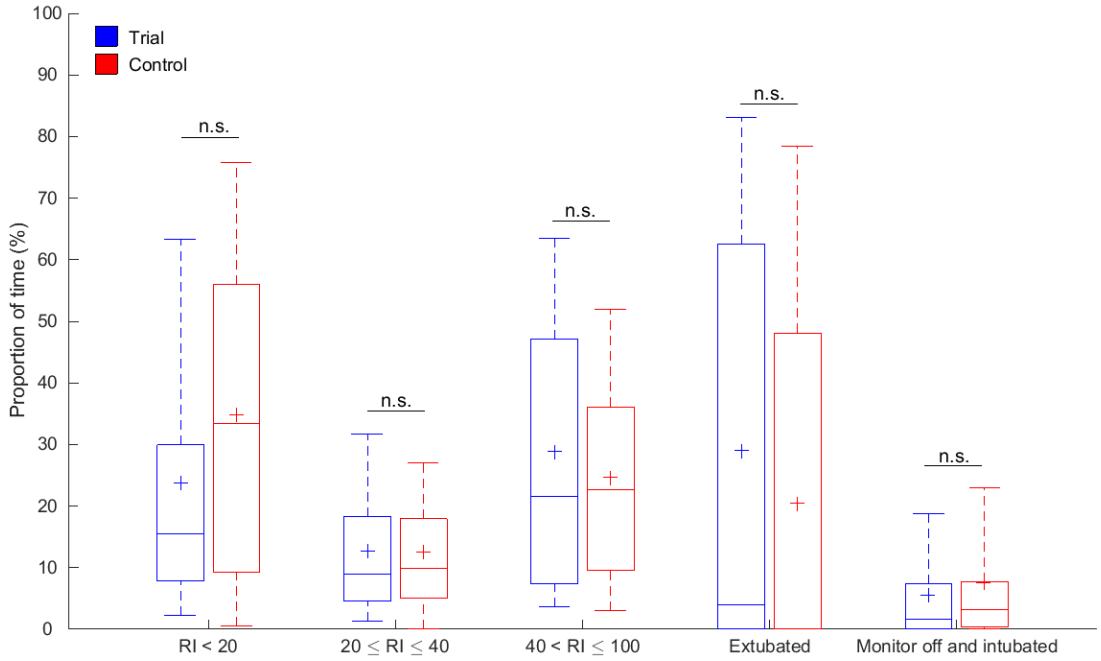


Figure 12: Proportions of time the patients spent on each RI category or extubated during the intervention period. Trial group values in left-hand side in blue and control group values in right-hand side in red. Cross-line within box indicates median and + value is mean while the upper box limit is 3rd quartile and lower box limit is 1st quartile. Whiskers represent the range between 9-91%. No statistical significance is denoted with n.s..

5.3 RASS scores

5.3.1 Day and night analysis

Day and night analysis of the RASS scores (Figure 13) shows that the RASS scores were lower during the night time. The median RASS score during the day time was –2 (n = 821) and during the night time was –3 (n = 509). This result was statistically significant (Wilcoxon rank-sum test $p < 0.01$). Both groups had similar proportions of RASS scores taken during the day time and during the night time (see table 14).

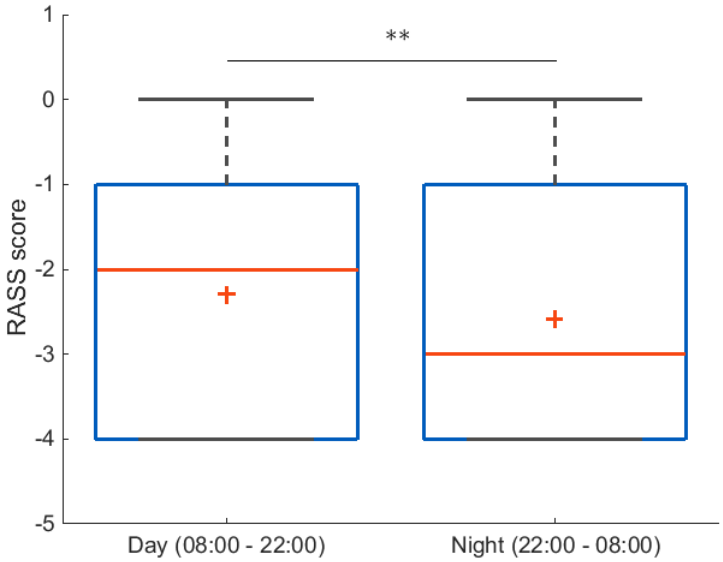


Figure 13: Differences in day time and night time RASS scores. A p-value < 0.01 is denoted with **.

Table 14: Proportions day and night RASS scores for both groups.

	Day (08:00-22:00)	**Night (22:00-08:00)**
Trial group	64.8%	35.2%
Control group	59.4%	40.6%

5.3.2 Time-of-day RASS evolution

Figure 14 shows the evolution of RASS scores in given times of day during the intervention period. According to the figure, there were no differences between the groups in number of patients with RASS ≤ -4 during the second day, when all patients had been recruited to the study.

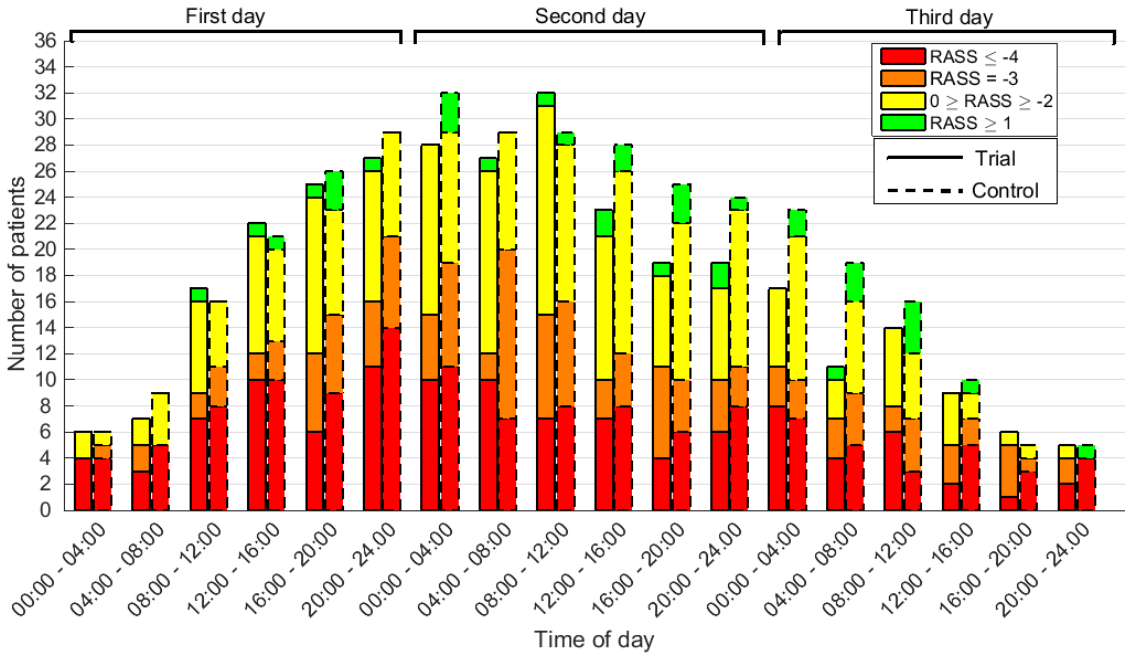
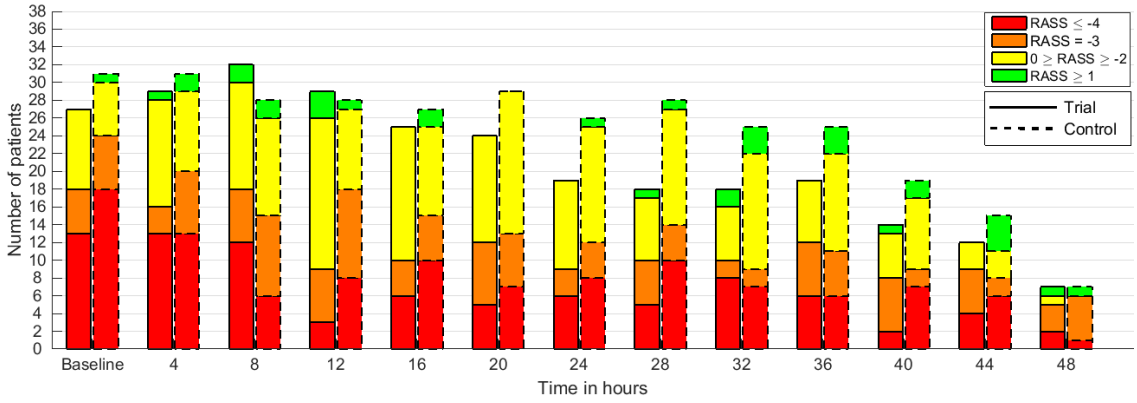


Figure 14: Time of day RASS evolution. A RASS sample was selected from each patient every four hours after the start of monitoring. A maximum of 48 hours of data were used for all patients. Number of patients monitored with RASS varies between timepoints because RASS scoring could not be performed every hour by the clinical staff or the patient was asleep. Left-hand side values belong to the trial group (solid bar lines) and right-hand side values to the control group (bars outlined with '- -')

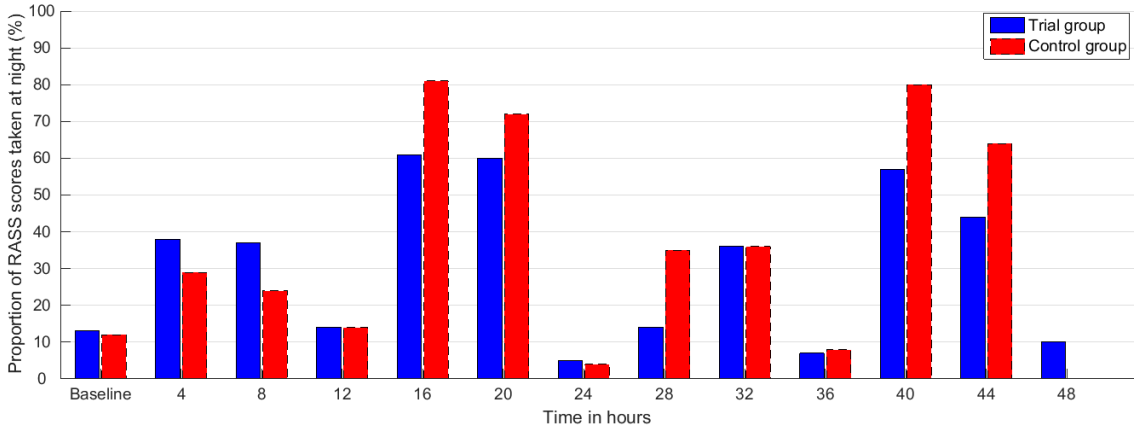
5.3.3 RASS evolution as a function of time elapsed from start of monitoring

Figure 15a presents the evolution of RASS scores during the intervention period as a function of time elapsed from start of monitoring. In figure 15b, the proportion of RASS scores taken at night time in each time point are shown. The RASS scores at baseline are the first RASS scores after the start of monitoring. The figure does not indicate that patients in the trial group would have less incidences of deep sedation (RASS $\leq$ –4). The number of deeply sedated patients clearly falls in both groups, but there is no distinct difference between the trial and control group.

Figure 15b shows that the number of RASS scores taken at night time varied between the time points. However, both groups had virtually the same number of RASS scores taken at night time in each time point.



(a)



(b)

Figure 15: (a) Evolution of RASS scores during the intervention period as a function of time elapsed from start of monitoring. Number of patients monitored with RASS varies between timepoints because RASS scoring could not be performed every hour by the clinical staff or patient was asleep. Left-hand side values belong to the trial group (solid bar lines) and right-hand side values to the control group (bars outlined with '- -'). (b) Proportion of RASS scores taken at night in given time points.

5.3.4 Time to reach first RASS > -4

The Kaplan–Meier curve (figure 16) shows the time from start of monitoring to first RASS > –4. The figure does not indicate that there would be any difference between the groups in terms of time to reach RASS > –4 (Logrank test $p = 0.677$).

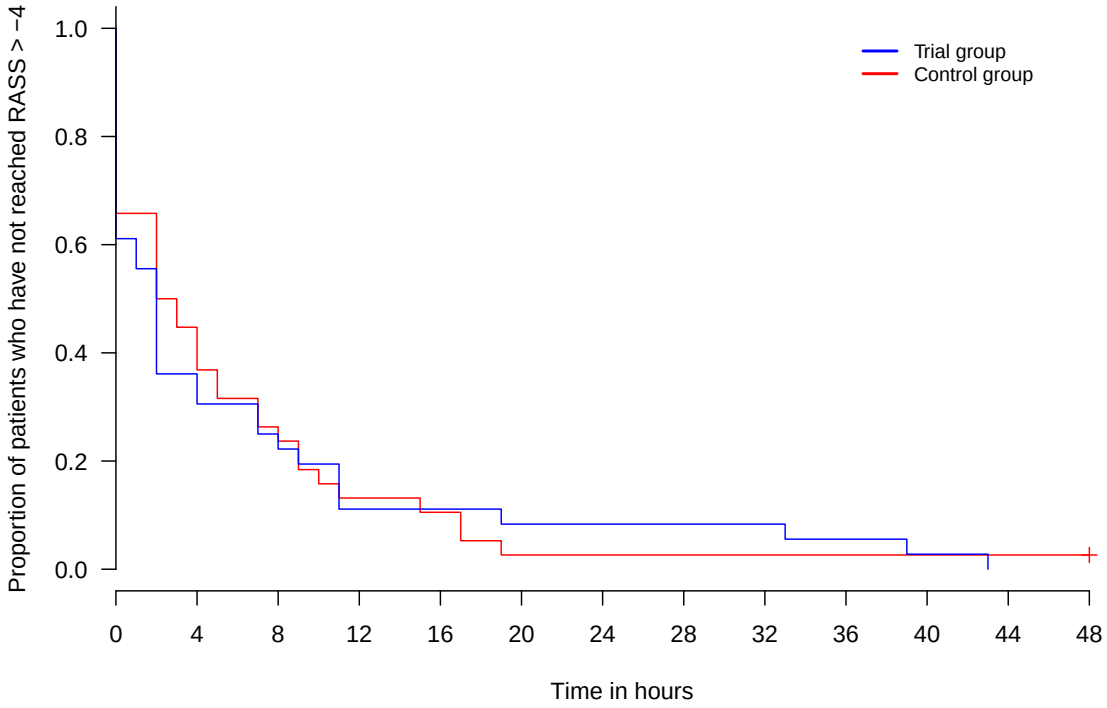


Figure 16: Kaplan–Meier curve for first RASS > –4. Censored events are marked with cross. One patient in control group never reached RASS score > –4, and therefore was censored. A total of 27 patients (n = 14 in the trial group and n = 13 in the control group) had first RASS ≥ –4 after the start of monitoring, hence were censored.

5.3.5 Proportions of RASS scores after the intervention period

The box plot (Figure 17) shows the distribution of RASS scores among the patients after the intervention period. According to the figure, both groups had similar proportions of RASS scores. The primary category of interest was the proportion of RASS scores ≤-4, and the other categories were plotted out of curiosity.

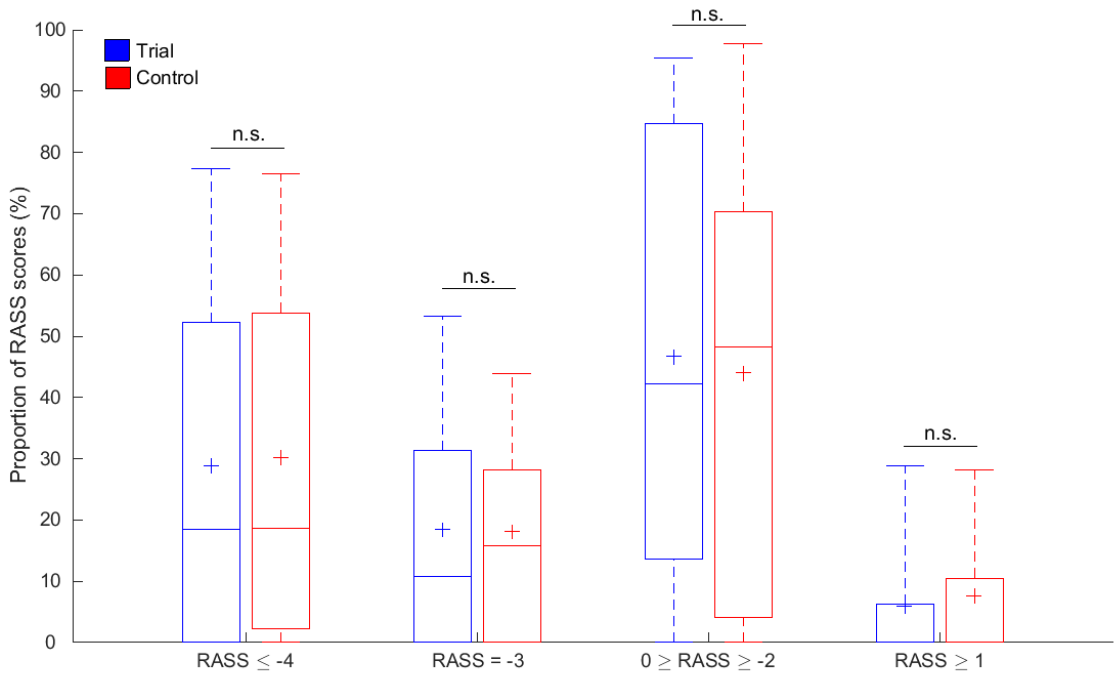


Figure 17: Distribution of RASS scores among the patients after the intervention period. Trial group values in left-hand side in blue and control group values in right-hand side marked with red.

5.4 Summary of results for RI and RASS analyses

Table 15 summarizes the results of the RI and RASS analyses with hypothesis testing. In conclusion, none of the measured parameters showed statistically significant results with p-value < 0.05.

Table 15: Intervention period variables

	Group		**Hypothesis test**
Intervention period variables	Trial	Control	P-value
Proportion of time during the intervention period with RI < 20, median (1st, 3rd quartile; min-max)	15.6% (8.4%, 29.6%; 0.9% - 86.9%)	33.4% (10.4%, 53.8%; 0.0% - 85.6%)	0.077^{W}
Proportion of incidences with RASS score of –4 or –5 during the intervention period, median (1st, 3rd quartile; min-max)	11.9% (0.0%, 35.4%; 0.0% - 93.5%)	16.9% (2.7%, 36.2 %; 0.0% - 100.0%)	0.698^{W}
Time (in hrs) from start of monitoring to first RI ≥ 20, median (1st, 3rd quartile; min-max)	0.09 (0, 2.22, 0.00 – 7.84)	0.33 (0.00, 3.08, 0 – NA)	0.428^{L}
Time (in hrs) from start of monitoring to first RASS > -4, median (1st, 3rd quartile; min-max)	2 (0, 8, 0 - 43)	3 (0, 8, 0 - NA)	0.677^{L}

^{W} Wilcoxon rank-sum test p-value
^{L} Logrank test p-value

5.5 Outcome measures

5.5.1 Time to first extubation

The Kaplan–Meier curve (Figure 18) shows the time to first extubation for both groups. The follow-up was done until 7 days (>168 hours) had elapsed from the start of monitoring. The baseline was defined as the time when monitoring starts. According to the baseline characteristics (Table 11), the time from intubation to start of monitoring was similar in both groups, and therefore, the groups are comparable.

The curve indicates that the patients in trial group were extubated faster during the intervention period, but the difference between the groups evens out after 72 hours, and the result was not statistically significant (Logrank test $p = 0.517$).

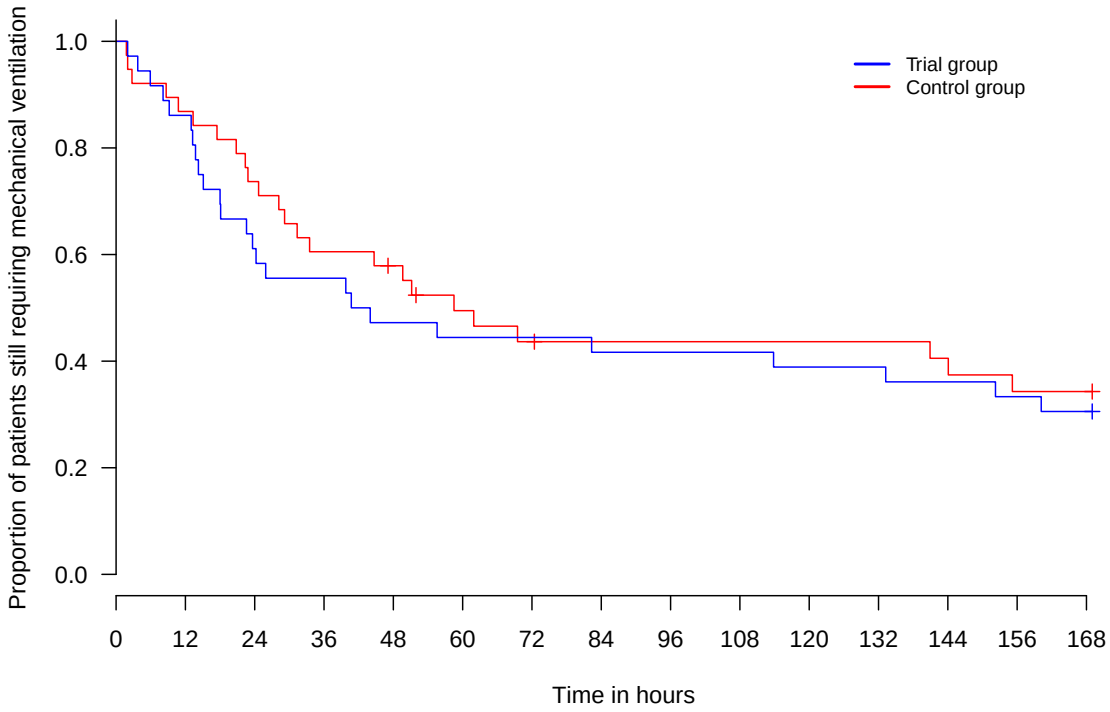


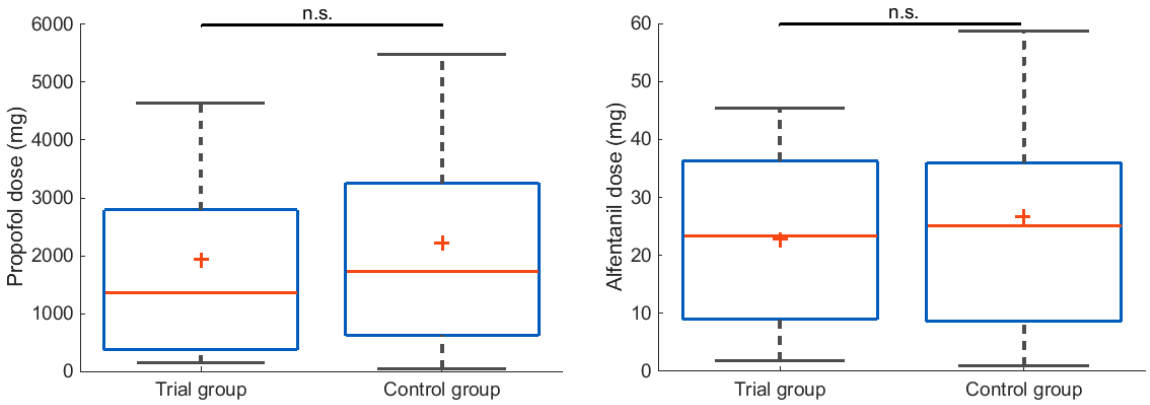
Figure 18: Kaplan-Meier curve of time to first extubation. Censored events are marked with cross. Three patients in control group were censored during the follow-up time due to the following reasons: transferred to another ICU (n = 2) or self-extubated (n = 1). For a total of 22 patients (n = 11 in the trial group and n = 11 in the control group), the extubation time was over seven days (> 168 hours), and therefore they were censored.

5.5.2 Total drug doses during the intervention period

Figure 19 shows the distribution of administered drug doses for propofol and alfentanil during the intervention period. All patients in trial group and 36 patients in control group received propofol. Additionally, 33 patients in trial group and 37 patients in control group received alfentanil.

For the purpose of reporting, 1 mg of midazolam was considered equivalent to 10 mg propofol. Four patients in trial group and four patients in control group received midazolam. In addition, 1 mg of morphine was considered equivalent to 100 μg of alfentanil. Eight patients in trial group and seven patients in control group received morphine.

Minor differences can be seen between the groups in terms of drug doses. The median value of total propofol dose during the intervention period was smaller in trial group (1365 mg vs. 1730 mg). In addition, the median value of total alfentanil dose during the intervention period was also smaller in trial group (23.4 mg vs. 25.2 mg). However, these results were not significant (Wilcoxon rank-sum test p=0.638 for propofol and p=0.698 for alfentanil).



(a) Propofol

(b) Alfentanil

Figure 19: Distribution of total propofol and alfentanil doses during intervention period.

5.5.3 Summary of patient outcome results

Table 16 shows results of patient outcomes after the intervention period. No statistically significant results were found. In Table 17, the time to first extubation and patient outcomes after 7 days are summarized. No significant differences were found between the groups.

Table 16: Patient outcomes after the intervention period.

	Group		**Hypothesis test**
Outcome variables	Trial	Control	p-value
Total propofol dose (in mg) during the intervention period, median (1st, 3rd quartile; min-max) †	1365 (380, 2790; 50 - 8710)	1730 (620, 3260; 0 - 8630)	0.638^{W}
Total alfentanil dose (in mg) during the intervention period, median (1st, 3rd quartile; min-max) ‡	23.4 (9.0, 36.4; 0.0 - 51.5)	25.2 (8.6, 36.0; 0.3 - 87.0)	0.681^{W}
Proportion of patients undergoing sedation hold during the intervention period, n (%)	30/36 (83.3%)	33/38 (86.8%)	0.751^{F}
Extubated, n(%)	17/36 (47.2%)	17/36 (44.7%)	1.000^{F}
Died, n(%)	0 (0%)	0 (0%)	1.000^{F}
Number of patients with any pre-defined adverse events during the intervention period, n (%)	7/36 (19.4%)	5/38 (13.1%)	0.538^{F}
Unplanned extubation	0	1	1.000^{F}
Unplanned removal of vascular catheter	0	0	1.000^{F}
Unplanned removal of nasogastric or other enteral tube	1	0	0.487^{F}
Unplanned removal of other drain or device	0	0	1.000^{F}
Episode of myocardial ischaemia	0	0	1.000^{F}
Myocardial infarction	0	0	1.000^{F}
Episode of agitation requiring bolus treatment (rescue medication)	6	4	0.510^{F}

^{W} Wilcoxon rank-sum test p-value

^{F} Fisher's exact test p-value

† 1 mg of midazolam was considered equivalent to 10 mg of propofol. Eight patients received midazolam (n = 4 in the trial group and n = 4 in the control group)

‡ 1 mg of morphine was considered equivalent to 100 μg alfentanil. 15 patients received morphine (n = 8 in the trial group and n = 7 in the control group)

Table 17: Time to first extubation and patient outcomes after 7 days.

	Group		**Hypothesis test**
Outcome variables	Trial	Control	p-value
Time (in hrs) from start of monitoring to first extubation, median (1st, 3rd quartile; min-max)	42.4 (14.7, 168.0; 2.0 - NA)	54.8 (22.8, 168.0; 1.78 - NA)	0.517^{L}
Died, n (%)	1/36 (2.8%)	3/38 (7.9%)	0.615^{F}
Discharged from ICU, n (%)	20/36 (55.6%)	18/38 (47.4%)	0.497^{F}
Still in ICU, n (%)	15/36 (41.7%)	15/38 (41.7)	1.000^{F}
Still in ICU receiving mechanical ventilation, n	15	15	1.000^{F}
Transferred	0/36 (0%)	2/38 (2%)	0.494^{F}

^{L} Logrank test p-value
^{F} Fisher's exact test p-value

5.6 Secondary analysis

A total number of 39 patients (n = 19 in the trial group and n = 20 in the control group) had baseline RI < 20. Baseline RASS was ≤ -4 for 31 patients (n = 13 in the trial group and n = 18 in the control group). A total of 18 patients (n =8 in the trial group and n = 10 in the control group) had baseline RI < 20 and RASS $\leq$ -4.

In the subgroup, where the baseline RI < 20 (see table 18), the secondary analysis reveals a statistical difference between the groups in terms of proportion of time spent with RI < 20 (median of 16.1% in trial group vs. median of 51.4% in control group, $p = 0.016$). Additionally, the trial group patients received significantly less alfentanil during the intervention period (median of 21.2 mg in the trial group vs. median of 32.2 mg in the control group, $p = 0.011$). The median total propofol dose was $> 50\%$ smaller in trial group (1090 mg in the trial group vs. 2380 mg in the control group), but this result was not statistically significant ($p = 0.140$).

Table 19 does not indicate that the groups were significantly different with respect to any measured parameter in the subgroup where the baseline RASS score was $\leq$-4.

Table 18: Patients with baseline RI < 20.

	Group		**Hypothesis test**
Intervention period variables	Trial	Control	p-value
n	19	20	NA
Proportion of time during the intervention period with RI < 20, median (1st, 3rd quartile; min-max)	16.1% (11.4%, 45.4%; 2.4% - 86.9%)	51.4% (33.4%, 71.6%; 4.8% - 85.6%)	0.016^W
RASS score at baseline, median (1st, 3rd quartile; min-max)	-4 (-4, -3; min=-5, max=-2)	-4 (-4, -3; min=-5, max=0)	0.759^W
Proportion of incidences with RASS score of –4 or –5 during intervention period, median (1st, 3rd quartile; min-max)	46.6% (15.4%, 65.4%; 0.0% - 93.6%)	35.3% (9.6%, 72.3%; 0.0% - 1.0%)	0.632^W
Total propofol dose (in mg) during intervention period, median (1st, 3rd quartile; min-max) †	1090 (375, 2965; 100 - 7290)	2380 (1510, 3730; 60 - 8630)	0.140^W
Total alfentanil dose (in mg) during intervention period, median (1st, 3rd quartile; min-max) ‡	21.2 (8.5, 27.9; 0.0 - 49.0)	32.3 (23.3, 49.8; 1.0 - 87.0)	0.011^W

^{W} Wilcoxon rank-sum test p-value

† 1 mg of midazolam was considered equivalent to 10 mg of propofol. Five patients received midazolam (n = 2 in the trial group and n = 3 in the control group)

‡ 1 mg of morphine was considered equivalent to 100 μg alfentanil. Four patients received morphine (n = 2 in the trial group and n = 2 in the control group)

Table 19: Patients with baseline RASS≤-4

	Group		Hypothesis test
Intervention period variables	Trial	Control	p-value
n	13	18	NA
Proportion of time during the intervention period with RI value <20, median (1st, 3rd quartile; min-max)	21.4% (13.6%, 35.5%; 1.9% - 86.9%)	33.8% (13.6%, 56.3%; 0.0% - 85.6%)	0.412^{W}
RI value at baseline, median (1st, 3rd quartile; min-max)	4 (0, 64; 0 - 96)	1 (0, 31; 0 - 100)	0.424^{W}
Proportion of incidences with RASS score of -4 or -5 during the intervention period, median (1st, 3rd quartile; min-max)	32.1% (0.0%, 59.0%; 0.0% - 93.6%)	13.8% (0.0%, 30.6%; 0.0% - 77.8%)	0.555^{W}
† Total propofol dose (in mg) during intervention period, median (1st, 3rd quartile; min-max)	1520 (385, 2623; 100 - 8710)	2155 (620, 4140; 0 - 8630)	0.589^{W}
‡ Total alfentanil dose (in mg) during intervention period, median (1st, 3rd quartile; min-max)	22.5 (11.4, 35.1; 0.0 - 51.5)	21.8 (19.0, 36.5; 1.0 - 87.0)	0.810^{W}

^{W} Wilcoxon rank-sum test p-value
† 1 mg of midazolam was considered equivalent to 10 mg of propofol. Three patients received midazolam (n=1 in trial group and n=2 in control group)
‡ 1 mg of morphine was considered equivalent to 100 μg alfentanil. Five patients received morphine (n=3 in trial group and n=4 in control group)

Because the Table 18 revealed statistically significant results in the subgroup where the baseline RI < 20, the patients were also studied in terms of differences based on baseline RI (see Table 20). This was done in order to find out whether or not patients with baseline RI < 20 had common factors that could be identified in view of inclusion criteria for the pivotal trial. However, no common factors were found as the Table 20 shows that patients who had RI<20 at baseline were similar to patients with baseline RI ≥ 20 with respect to all compared variables.

Table 20: Comparison of patients' baseline characteristics with respect to baseline RI.

	Baseline RI		**Hypothesis test**
Variables	Baseline RI < 20	Baseline RI ≥ 20	p-value
Number of patients	39	35	NA
Age, median (1st, 3rd quartile; min-max)	59 (43, 71; 25 - 85)	62 (45, 69; 35 -79)	0.693^{W}
Sex, male/female	22/17	25/10	0.160^{F}
RASS score at baseline, median (1st, 3rd quartile; min-max)	-3 (-4, -2: min=-5, max=1)	-3 (-4, -2: min=-5, max=2)	0.764^{W}
Total propofol dose (mg) prior to monitoring, median (1st, 3rd quartile; min-max)†	370 (160, 575; 0 - 7200)	350 (150, 790; 0 - 1690)	0.867^{W}
Total alfentanil dose (mg) prior to monitoring, median (1st, 3rd quartile; min-max) ‡	2.0 (0.6, 7.0; 0.0 - 17.5)	2.5 (1, 6.3; 0.0 - 15.5)	0.638^{W}

^{W} Wilcoxon rank-sum test p-value
^{F} Fisher's exact test p-value
† 1 mg of midazolam was considered equivalent to 10 mg of propofol. One patient with low baseline RI and one patient with high baseline RI received midazolam prior to monitoring.
‡ 1 mg of morphine was considered equivalent to 100 μg alfentanil. One patient with low baseline RI and none of the patients with high baseline RI received morphine prior to monitoring.

6 Discussion

6.1 Main findings

As this was a pilot trial, statistically significant results were not expected to be found. However, the results of the primary analysis showed a trend in favor of trial group in terms of proportion of time spent on RI < 20, time to reach first RI ≥ 20, and time to first extubation during the intervention period. Nonetheless, without knowing the decision-making process of the practicioners that led to extubation, it is impossible to determine whether the use of RI actually led to a possibly shorter extubation times in trial group.

There were no differences between the groups in terms of pre-defined adverse events. This result may verify that using RI according to the study protocol is safe. However, the sample size was small and the sedation related adverse events do not occur frequently, so therefore studies with larger sample size will be needed to validate the safety of the use.

The time evolution analysis of RI revealed that RI values in trial group were notably higher compared to control group values between baseline and 24 hours after start of monitoring. This result indicates that the use of RI monitoring is most beneficial during this time period. The clinical implication of the finding is that RI can help to guide the caregiving staff to optimally sedate the patients early after intubation when patients require high doses of sedatives.

No trend was found indicating that there would be any difference between the groups in terms of RASS scores, and in times to reach first RASS > -4 during the intervention period. This raises the question, why the trend could be seen in RI values but not in RASS scores? The reason might be that the purpose of RASS and RI are not identical, thus they cannot be compared. RASS relies on measuring the effects of fixed stimuli in a discrete time point, while RI is a continuous retrospective measurement of the patient's responsiveness to ambient and internal stimuli.

The Kaplan–Meier curve (Figure 18) showed that some patients were extubated very early after the start of monitoring. This means that these patients in trial group were affected by the study protocol for only a little of time as RI monitoring was removed from patients who were extubated. It can be assumed that these patients also received only small doses of sedative drugs during the intervention period, because they were quickly removed from mechanical ventilator. In conclusion, if the patient only received small amounts of sedatives and spent only a short time in mechanical ventilation, he or she was likely unaffected by the study protocol, because the nurses did not have enough time to make decisions based on RI, and there were not a high doses of sedatives from where to reduce.

6.2 A comparison to similar study done with BIS

The results of a similar study by Olson and colleagues [73] performed with BIS-augmented sedation monitoring were briefly presented in Chapter 3.1.1. In the study, Olson and the research team reported that using BIS as an adjunct to RSS (n =

32) compared to only using RSS (n = 35), decreased the amounts of administered sedatives significantly (93.5 ml vs. 157.8 ml, $p < 0.015$). This raises the question why in the study by Olson and colleagues the research team was able to have statistically significant results in terms of drug doses and this study could not? This can be expained with the following reasons:

1. The studied time period was different. Olson et al. used 12-hour measurement period, and in this study the maximum time period of RI monitoring was 48 hours. The sedation monitoring might be most beneficial during the first 12 (or maybe 24) hours after intubation, because during this time period the patients need more sedatives. After this critical period, the sedation monitoring with RI or BIS might be insignificant.

2. The sedation management and the quality of sedation management varies between hospitals. In this study, the control group patients were treated according to the defined best practice. The article published by Olson et al. does not reveal, in detail, how the sedation management was performed in patients who did not receive BIS agumented sedation monitoring. They only disclose that these patients were monitored with RSS and the nurses were instructed to adjust the sedation infusion to achieve RSS score of 4 (patient asleep, brisk response to glabellar tap or loud auditory stimulus). A multicenter trial would be needed in order to validly compare the study protocols and performances of the monitors.

6.3 Secondary analysis

The secondary analysis revealed that the effects of RI monitoring were more visible in the subgroup where baseline RI < 20. These patients spent more time with RI < 20 during the intervention period, and therefore, were more affected by the study protocol. As a result, the patients in trial group received significantly less alfentanil during the intervention period. Additionally, a clear trend was seen showing that the patients in trial group received less propofol ($p = 0.014$).

In the subgroup where baseline RASS ≤ -4, no differences were found between the groups with respect to any variable. The reason for this result might be that the administration of sedatives was not primarily controlled with RASS.

It must be noted that the baseline characteristics of these two subgroups were not analyzed. However, as these subgroups were taken from the same patient population, the baseline characteristics were expected to be similar between the trial and control group.

6.4 Recommendations for pivotal trial

A flowchart of the recommended study protocol for the pivotal trial is showed in Figure 20, and proposed treatment algorithm for trial group patients in Figure 21. Few modifications are proposed to the protocol tested in this study.

Firstly, based on the result that the effects of RI monitoring are significant among patients with baseline RI < 20, the primary analysis should be performed only for this subgroup of patients. The patients who had baseline RI < 20 were more likely to spend more time with low RI values, and therefore, were more affected by the tested study protocol. The secondary analysis should be performed for all patients.

Secondly, based on the significant findings of this study, the main objective in the pivotal trial should be to test whether RI-augmented group, compared to control group, will receive significantly less sedatives and analgesics during the monitoring period without excess adverse events. Therefore, the total sedative and analgesic drug doses and the number of adverse events during the monitoring period should be called the primary measures of the study. As the primary analysis and secondary analysis of this thesis did not find any difference in terms of proportions of deep sedation (RASS ≤ -4) between the groups, the primary measure in the pivotal trial should not be the number of deep sedation incidences. Additionally, as RI does not have a valid reference, and no study has yet done which would prove that RI is a valid measure of sedation status, the primary measure of the pivotal trial should not be the proportion of time spent on low RI levels. The secondary measures of the pivotal trial should be the other patient outcomes explored in this pilot trial with high priority in time to reach first extubation. The follow-up for the secondary measures should be continued until at least 7 days has elapsed since ICU admission.

Thirdly, the time evolution analyses of RI showed that RI monitoring was most effective between baseline and 24 hours. Therefore, the RI monitoring period should be from start of monitoring until 24 hours has elapsed from ICU admission. This study also showed that time from intubation to start of monitoring varied between 1.7 hours to 12 hours. The study protocol of the pivotal trial needs emphasize the importance of starting the RI monitoring as early as possible after the intubation. This way the nurses will receive RI-augmented sedation guidance during the period which is the most critical in sedation management.

Lastly, as this study showed that RI monitoring does not increase the number of sedation related adverse events, the patients in trial group should be treated with sedation-reduction algorithm that could lead to more significant results. The algorithm is described in figure 21. The objective of the algorithm is to reach and maintain RI ≥ 40, and a state in which the patient responds to verbal stimulation. Increasing the RI threshold value to 40 can be expected to be safe, as the sedation is reduced also based on whether or not the patient responds to verbal stimulation. Additionally, the algorithm has more structured approach to use the RI monitoring by setting rules when to observe the RI value. This way the nurses cannot neglect the RI monitor, and are obligated to use RI monitoring when treating the patients. The waiting time after patient responds to verbal stimulation was set to 60 minutes so that the patient would not be disturbed frequently. All cases in which the described

algorithm cannot be followed should be reported.

The patients in the control group should be treated with the current hospital policy, not with best practice. This way the study would reflect the benefits of RI-augmented sedation monitoring in the most true setting.

The pivotal trial should aim to provide valid scientific evidence to support the claim that the use of RI as an adjunct other sedation practices may be associated with a reduction in primary sedative use during early intensive care.

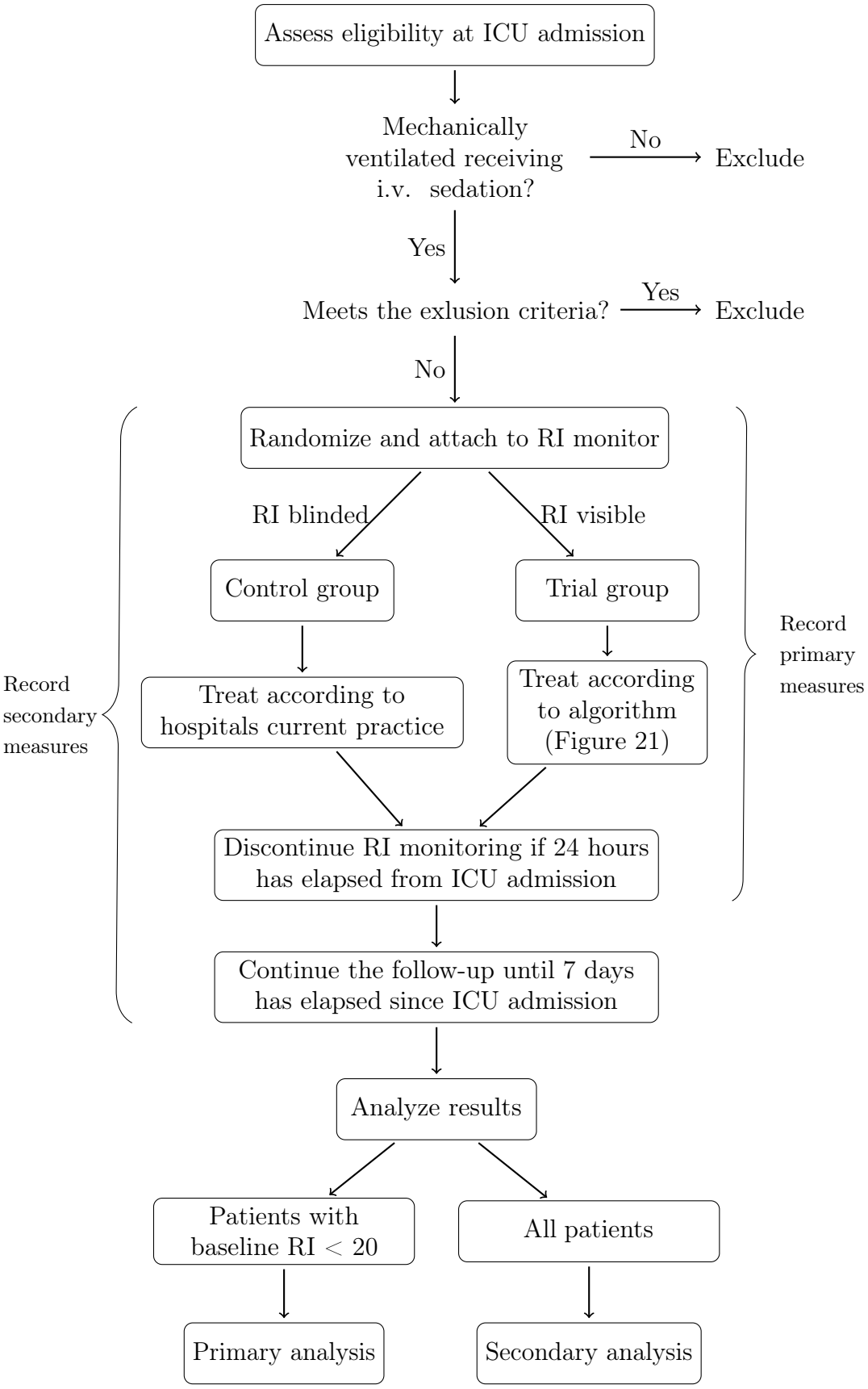


Figure 20: Flowchart of the proposed study protocol for the pivotal trial.

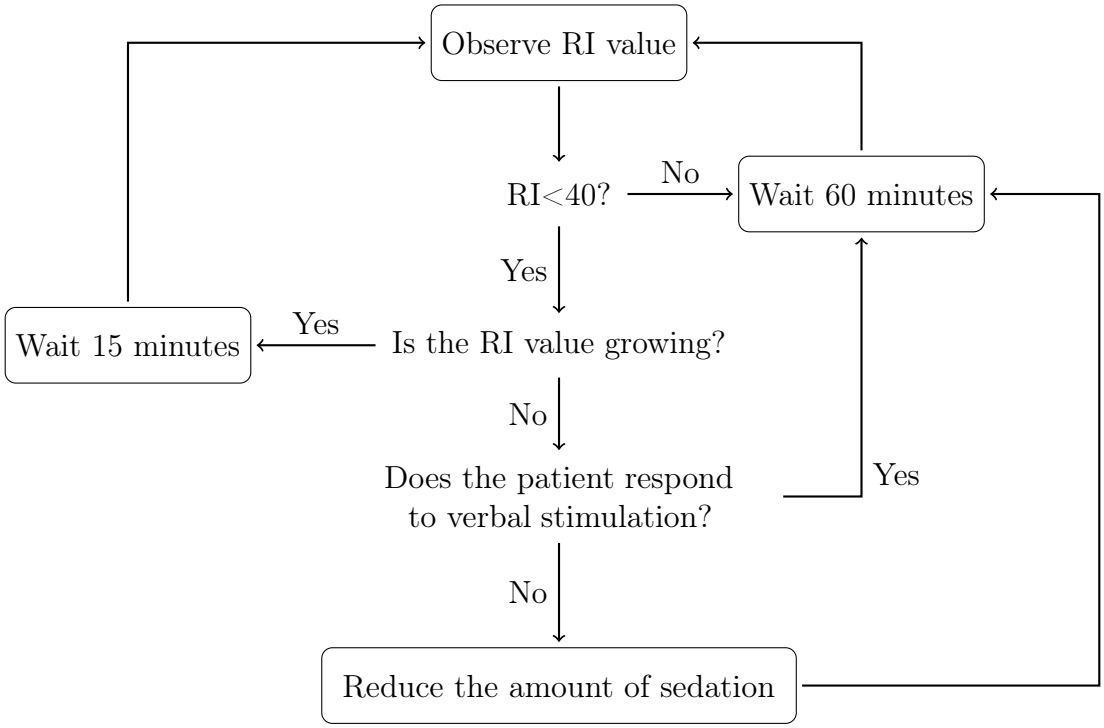


Figure 21: Treatment algorithm for trial-group patients.

7 Conclusion

In this Thesis, the objective was to assess the effectiveness and safety of continuous Responsiveness Index-augmented sedation monitoring during early ICU care as a nurse decision-support tool. In conclusion, the RI monitoring seemed to guide the administration of sedatives and analgesics as RI values showed to be higher in trial group. In addition, there were no excess adverse events, and therefore, the protocol was safe for the investigated patients.

The results were promising. Multiple outcome measures indicated that RI monitoring might lead to better sedation outcomes without excess adverse events. However, as this study had a small sample size, the results in the primary analysis were not statistically significant, but a trend was found.

This study proved that the protocol could be used in the pivotal trial with few modifications. It also supports the concept of RI-augmented sedation monitoring as a tool to safely reduce the amount of sedative drugs administered to mechanically ventilated patients. Additional studies with larger sample sizes will be needed to fully determine the benefits and safety of using fEMG based monitors in sedation management.

References

[1] A. Batchelor and P. Nightingale, "General principles of intensive care management," in *ABC of Intensive Care* (G. Nimmo and M. Singer, eds.), pp. 14–19, Wiley, 2nd ed., 2011.

[2] S. Laha, N. Arora, and J. Goodall, *A Beginner's Guide to Intensive Care Medicine*. Radcliffe Publishing, 2010.

[3] C. Sessler and K. Muzevinch, "Use of sedatives, analgensics, and neuromuscular blockers," in *Critical Care Medicine: Principles of Diagnosis and Management in the Adult* (J. Parillo and P. Dellinger, eds.), pp. 255–271, Elsevier Health Sciences, 4th ed., 2013.

[4] NHS, "Intensive care." http://www.nhs.uk/conditions/intensive-care/pages/introduction.aspx. Accessed January 6, 2015.

[5] D. O'Callaghan, P. Jayia, E. Vaughan-Huxley, M. Gribbon, M. Templeton, J. Skipworth, and A. Gordon, "An observational study to determine the effect of delayed admission to the intensive care unit on patient outcome," *Critical Care*, 2012. R173.

[6] H. Wunsch, D. Angus, D. Harrison, and W. Linde-Zwirble, "Comparison of medical admissions to intensive care units in the united states and united kingdom," *American Journal of Respiratory and Critical Care Medicine*, vol. 183, pp. 1666–1673, 2011.

[7] Society of Critical Care Medicine, "Critical care statistics." www.sccm.org/Communications/Pages/CriticalCareStats.aspx. Accessed January 8, 2015.

[8] T. Huynh, E. Kleerup, J. Wiley, T. Savitsky, D. Guse, B. Garber, and N. Wenger, "The frequency and cost of treatment perceived to be futile in critical care," *JAMA Internal Medicine*, vol. 173, pp. 1887–1894, 2013.

[9] Y. Arabi, S. Venkatesh, S. Haddad, A. Al Shimemeri, and S. Al Malik, "A prospective study of prolonged stay in the intensive care unit: predictors and impact on resource utilization," *International Journal for Quality in Health Care*, vol. 14, pp. 403–410, 2002.

[10] C. Hughes, S. McGrane, and P. Pandharipande, "Sedation in the intensive care setting," *Clinical Pharmacology*, vol. 4, pp. 53–63, 2012.

[11] N. Safdar, C. Dezfulian, H. Collard, and S. Saint, "Clinical and economic consequences of ventilator-associated pneumonia: a systematic review," *Critical Care Medicine*, vol. 33, pp. 2184–2193, 2005.

[12] A. Leon, L. Davis, and H. Kraemer, "The role and interpretation of pilot studies in clinical research," *Journal of Psychiatric Research*, vol. 45, pp. 626–629, 2010.

[13] J. Loscalzo, "Pilot trials in clinical research," *Circulation*, vol. 119, pp. 1694–1696, 2010.

[14] J. Barr, G. Fraser, K. Puntillo, W. Ely, C. Gelinas, J. Dasta, J. Davidson, *et al.*, "Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit," *Critical Care Medicine*, vol. 41, pp. 263–306, 2013.

[15] J. Barr and P. Pandharipande, "The pain, agitation, and delirium care bundle: synergistic benefits of implementing the 2013 pain, agitation, and delirium guidelines in an integrated and interdisciplinary fashion," *Critical Care Medicine*, vol. 41, pp. 99–115, 2013.

[16] M. Tanios, M. de Wit, S. Epstein, and J. Devlin, "Perceived barriers to the use of sedation protocols and daily sedation interruption: a multidisciplinary survey," *Journal of Critical Care*, vol. 24, pp. 66–73, 2009.

[17] J. Slomka, L. Hoffman-Hogg, L. Mion, N. Bair, M. Bobek, and A. Arroliga, "Influence of clinicians' values and perceptions on use of clinical practice guidelines for sedation and neuromuscular blockade in patients receiving mechanical ventilation," *American Journal of Critical Care*, vol. 9, pp. 412–418, 2000.

[18] A. Piriyapatsom, E. Bittner, J. Hines, and U. Schmidt, "Sedation and paralysis," *Respiratory Care*, vol. 58, p. 1024–1037, 2013.

[19] K. Rowe and S. Fletcher, "Sedation in the intensive care unit," *British Journal of Anaesthesia*, vol. 8, pp. 50–55, 2008.

[20] B. Pun and J. Dunn, "The sedation of critically ill adults: Part 1: Assessment," *The American Journal of Nursing*, vol. 107, pp. 40–48, 2007.

[21] J. Nelson, D. Meier, E. Oei, D. Nierman, R. Senzel, P. Manfredi, S. Davis, *et al.*, "Self-reported symptom experience of critically ill cancer patients receiving intensive care," *Critical Care Medicine*, vol. 29, pp. 277–282, 2001.

[22] M. Reade and S. Finfer, "Sedation and delirium in the intensive care unit," *The New England Journal of Medicine*, vol. 370, pp. 444–454, 2014.

[23] Y. Skrobik, S. Ahern, M. Leblanc, Marquis, D. F Awissi, and B. Kavanagh, "Protocolized intensive care unit management of analgesia, sedation, and delirium improves analgesia and subsyndromal delirium rates," *Anesthesia and Analgesia*, vol. 111, pp. 451–463, 2010.

[24] T. Whitehouse, C. Snelson, M. Grouds, T. Willson, L. Tulloch, L. Linhartova, A. Shah, *et al.*, "Intensive care society review of best practice for analgesia and sedation in the critical care," 2014. http://www.ics.ac.uk/EasysiteWeb/getresource.axd?AssetID=2362. Accessed January 30, 2015.

[25] B. Shapiro, J. Warren, A. Egol, D. Greenbaum, J. Jacobi, S. Nasraway, R. Schein, *et al.*, "Practice parameters for intravenous analgesia and sedation for adult patients in the intensive care unit: an executive summary," *Critical Care Medicine*, vol. 23, pp. 1596–600, 1995.

[26] M. Ramsay, "How to use the Ramsay score to assess the level of ICU sedation." http://5jsnacc.umin.ac.jp/How%20to%20use%20the%20Ramsay%20Score%20to%20assess%20the%20level%20of%20ICU%20Sedation.htm. Accessed February 16, 2015.

[27] Institute for Healthcare Improvement, "How-to guide: Prevent harm from high-alert medications." http://www.ihi.org/resources/Pages/Tools/HowtoGuidePreventHarmfromHighAlertMedications.aspx. Accessed March 2, 2015.

[28] B. Pun and J. Dunn, "The sedation of critically ill adults: Part 1: Assessment," *The American Journal of Nursing*, vol. 107, pp. 40–49, 2007.

[29] K. McKeage and C. Perry, "Propofol: a review of its use in intensive care sedation of adults," *CNS Drugs*, vol. 17, pp. 235–272, 2003.

[30] J. Barr, "Propofol: a new drug for sedation in the intensive care unit," *International Anesthesiology Clinics*, vol. 33, pp. 131–154, 1995.

[31] AstraZeneca UK Limited, "Diprivan 1%." https://www.medicines.org.uk/emc/medicine/2275. Accessed January 30, 2015.

[32] A. Shafer, "Complications of sedation with midazolam in the intensive care unit and a comparison with other sedative regimens," *Critical Care Medicine*, vol. 26, pp. 947–56, 1998.

[33] Accord Healthcare Limited, "Midazolam 5 mg/ml solution for injection or infusion." http://www.medicines.org.uk/emc/medicine/25847. Accessed February 4, 2015.

[34] L. Mion, A. Minnick, R. Leipzig, C. Catrambone, and M. Johnson, "Patient-initiated device removal in intensive care units: a national prevalence study," *Critical Care Medicine*, vol. 35, pp. 2714–2720, 2007.

[35] D. Coursin and D. Coursin, "Survivors, beware of post-traumatic stress disorder: What shall we tell the men in black?," *Critical Care Medicine*, vol. 26, pp. 634–635, 1998.

[36] D. Jackson, C. Proudfoot, K. Cann, and T. Walsh, "The incidence of sub-optimal sedation in the ICU: a systematic review," *Critical Care*, vol. 13, pp. 1–14, 2009.

[37] D. Becker and D. Haas, "Management of complications during moderate and deep sedation: Respiratory and cardiovascular considerations," *Anesthesia Progress*, vol. 54, pp. 59–69, 2008.

[38] E. Vasilevskis, W. Ely, T. Speroff, B. Pun, L. Boehm, and R. Dittus, "Reducing iatrogenic risks: ICU-acquired delirium and weakness—crossing the quality chasm," *Chest*, vol. 138, pp. 1224–1233, 2010.

[39] T. Girard, "Living on the lighter side of sedation in the intensive care unit: Is there a psychological cost?," *Critical Care Medicine*, vol. 37, pp. 2654–2655, 2009.

[40] E. Whitaker, A. Mukherjee, T. Liu, B. Hong, and E. Heitmiller, "Introduction to moderate and deep sedation," in *Moderate and Deep Sedation in Clinical Practice* (R. Urman and A. Kaye, eds.), pp. 1–7, Cambridge University Press, 2012.

[41] D. Hogarth and J. Hall, "Management of sedation in mechanically ventilated patients," *Current Opinion in Critical Care*, vol. 10, pp. 40–46, 2004.

[42] C. Sessler, M. Gosnell, M. Grap, G. Brophy, P. O'Neal, K. Keane, E. Tesoro, and R. Elswick, "The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients," *American Journal of Respiratory and Critical Care Medicine*, vol. 166, pp. 1338–1344, 2002.

[43] S. Stawicki, "Sedation scales: Very useful, very underused," *OPUS 12 Scientist*, vol. 1, pp. 10–12, 2007.

[44] E. Ely, B. Truman, A. Shintani, J. Thomason, A. Wheeler, S. Gordon, J. Francis, *et al.*, "Monitoring sedation status over time in ICU patients: Reliability and validity of the Richmond Agitation-Sedation Scale (RASS)," *JAMA*, vol. 11, pp. 2983–2991, 2003.

[45] M. Ramsay, T. Savege, B. Simpson, and R. Goodwin, "Controlled sedation with alphaxalone-alphadolone," *British Medical Journal*, vol. 2, pp. 656–659, 1974.

[46] J. Payen, G. Chanques, J. Mantz, C. Hercule, I. Auriant, L. JC, M. Binhas, *et al.*, "Current practices in sedation and analgesia for mechanically ventilated critically ill patients," *Anesthesiology*, vol. 106, pp. 687–695, 2007.

[47] B. De Jonghe, D. Cook, C. Appere-De-Vecchi, G. Guyatt, M. Meade, and H. Outin, "Using and understanding sedation scoring systems: a systematic review," *Intensive Care Medicine*, vol. 26, pp. 275–285, 2000.

[48] M. Kollef, N. Levy, T. Ahrens, R. Schaiff, D. Prentice, and G. Sherman, "The use of continuous i.v. sedation is associated with prolongation of mechanical ventilation," *Chest*, vol. 114, pp. 541–548, 1998.

[49] J. Kress, A. Pohlman, M. O'Connor, and J. Hall, "Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation," *The New England Journal of Medicine*, vol. 342, pp. 1471–1477, 2000.

[50] J. Odom-Forren and D. Watson, *Practical Guide to Moderate Sedation/analgesia*. Elsevier Health Sciences, 2nd ed., 2005.

[51] T. Girard, J. Kress, B. Fuchs, J. Thomason, W. Schweickert, B. Pun, D. Taichman, *et al.*, "Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (awakening and breathing controlled trial): a randomised controlled trial," *Lancet*, vol. 371, pp. 126–134, 2008.

[52] S. Devabhakthuni, M. Armahizer, J. Dasta, and S. Kane-Gill, "Analgosedation: A paradigm shift in intensive care unit sedation practices," *The Annals of Pharmacotherapy*, vol. 46, pp. 530–540, 2012.

[53] C. Weinert, L. Chlan, and C. Gross, "Sedating critically ill patients: factors affecting nurses' delivery of sedative therapy," *American Journal of Critical Care*, vol. 10, pp. 156–165, 2001.

[54] T. Walsh, K. Everingham, F. Frame, P. Lapinlampi, M. Särkelä, K. Uutela, and H. Viertiö-Oja, "An evaluation of the validity and potential utility of facial electromyelogram responsiveness index for sedation monitoring in critically ill patients," *Journal of Critical Care*, vol. 29, pp. 886–887, 2014.

[55] S. Christensen, M. Johansen, C. Christensen, R. Jensen, and S. Lemenshow, "Comparison of charlson comorbidity index with SAPS and APACHE scores for prediction of mortality following intensive care," *Clinical Epidemiology*, vol. 3, pp. 203–211, 2011.

[56] N. Campbell, M. Tooley, and S. Willatts, "APACHE II scoring system on a general intensive care unit: audit of daily APACHE II scores and 6-month survival of 691 patients admitted to a general intensive care unit between may 1990 and december 1991," *Journal of the Royal Society of Medicine*, vol. 87, pp. 73–77, 1994.

[57] W. Knaus, E. Draper, D. Wagner, and J. Zimmerman, "APACHE II: A severity of disease classification system," *Critical Care Medicine*, vol. 13, pp. 818–829, 1985.

[58] J. Vincent and R. Moreno, "Clinical review: Scoring systems in the critically ill," *Critical Care*, vol. 14, pp. 1–9, 2010.

[59] F. Lopez Ferreira, D. Peres Bota, A. Bross, C. Melot, and J. Vincent, "Serial evaluation of the SOFA score to predict outcome in critically ill patients," *JAMA*, vol. 286, pp. 1754–1758, 2001.

[60] P. Marino and K. Sutin, *The ICU Book*. Lippincott Williams and Wilkins, 3rd ed., 2012.

[61] J. Vincent, "SOFA score: A keystone for grading multiple organ dysfunction," in *Anaesthesia, Pain, Intensive Care and Emergency* (A. Gullo, ed.), pp. 255–259, Springer Science and Business Media, 1st ed., 2008.

[62] M. Smith, N. Nitz, and S. Stuart, "Severity and comorbidity," in *Understanding Health Care Outcomes Research* (R. Kane, ed.), pp. 219–262, Jones and Bartlett Learning, 2nd ed., 2006.

[63] C. Kastner, J. Armitage, A. Kimble, J. Rawal, P. Cartner, and S. Venn, "The Charlson comorbidity score: a superior comorbidity assessment tool for the prostate cancer multidisciplinary meeting," *Prostate Cancer and Prostatic Diseases*, vol. 9, p. 270–274, 2006.

[64] D. Zuckerman, P. Brown, and S. Nissen, "Medical device recalls and the FDA approval process," *Archives of Internal Medicine*, vol. 171, pp. 1006–1011, 2011.

[65] A. Kaplan, D. Baim, J. Smith, D. Feigal, M. Simons, D. Jeffreys, T. Fogarty, *et al.*, "Medical device development from prototype to regulatory approval," *Circulation*, vol. 109, pp. 3068–3072, 2004.

[66] The US Food and Drug Administration, "K072286: 510(k) summary of Aspect Medical Systems BIS EEG VISTA Monitor System," 2007. http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K072286. Accessed March 3, 2015.

[67] The US Food and Drug Administration, "K051874: 510(k) summary of SedLine sedation monitor with Frontal PSI and SEDtrace EEG Electrode Set," 2005. http://www.accessdata.fda.gov/cdrh_docs/pdf5/K051874.pdf. Accessed March 18, 2015.

[68] The US Food and Drug Administration, "K061907: 510(k) summary of Datex-Ohmeda S5 Entropy Module, E-ENTROPY and accessories," 2008. http://www.accessdata.fda.gov/cdrh_docs/pdf6/K061907.pdf. Accessed March 18, 2015.

[69] T. Walsh, P. Ramsay, and R. Kinnunen, "Monitoring sedation in the intensive care unit: can "black boxes" help us?," *Intensive Care Medicine*, vol. 30, pp. 1511–1513, 2004.

[70] C. Sessler, M. Grap, and M. Ramsay, "Evaluating and monitoring analgesia and sedation in the intensive care unit," *Critical Care*, vol. 12, Suppl 3, S2, 2008.

[71] C. De Deyne, M. Struys, J. Decruyenaere, J. Creupelandt, E. Hoste, and F. Colardyn, "Use of continuous bispectral EEG monitoring to assess depth of sedation in ICU patients," *Intensive Care Medicine*, vol. 24, pp. 1294–1298, 1998.

[72] J. LeBlanc, J. Dasta, and S. Kane-Gill, "Role of the Bispectral Index in sedation monitoring in the ICU," *Critical Care*, vol. 40, pp. 490–500, 2006.

[73] D. Olson, S. Thoyre, E. Peterson, and C. Graffagnino, "A randomized evaluation of Bispectral Index -augmented sedation assessment in neurological patients," *Neurocritical Care*, vol. 11, pp. 20–27, 2009.

[74] A. Dahaba, "Different conditions that could result in the bispectral index indicating an incorrect hypnotic state," *Anesthesia and Analgesia*, vol. 101, pp. 765–773, 2005.

[75] D. Drover and H. Ortega, "Patient state index," *Best practice and Research Clinical Anaesthesiology*, vol. 20, pp. 121–128, 2006.

[76] G. Schneider, S. Heglmeier, J. Schneider, G. Tempel, and E. Kochs, "Patient state Index PSI measures depth of sedation in intensive care patients," *Intensive Care Medicine*, vol. 30, pp. 213–216, 2004.

[77] C. T, M. Ramsay, J. Rossmann, M. Beach, G. Griffiths, B. Meyrat, J. Barnes, *et al.*, "Evaluation of the SEDline to improve the safety and efficiency of conscious sedation," *Proceedings (Baylor University Medical Center)*, vol. 24, pp. 200–204, 2011.

[78] T. Walsh, P. Ramsay, T. Lapinlampi, M. Särkelä, H. Viertiö-Oja, and P. Merilainen, "An assessment of the validity of spectral entropy as a measure of sedation state in mechanically ventilated critically ill patients," *Intensive Care Medicine*, vol. 34, pp. 308–315, 2008.

[79] T. Walsh, T. Lapinlampi, P. Ramsay, M. Särkelä, K. Uutela, and H. Viertiö-Oja, "Responsiveness of the frontal EMG for monitoring the sedation state of critically ill patients," *British Journal of Anaesthesia*, vol. 107, pp. 710–718, 2011.

[80] J. Larsen, G. Berntson, K. Poehlmann, T. Ito, and J. Cacioppo, "The psychophysiology of emotion," in *Handbook of Emotions* (M. Lewis, J. Haviland-Jones, and L. Feldman Barrett, eds.), pp. 180–195, Guilford Press, 3rd ed., 2008.

[81] K. Pryor and R. Veselis, "Consciousness and cognition," in *Foundations of Anesthesia: Basic Sciences for Clinical Practice* (H. Hemmings and P. Hopkins, eds.), pp. 349–359, Elsevier Health Sciences, 2nd ed., 2006.

[82] U. Hess, "Facial EMG," in *Methods in Social Neuroscience* (E. Harmon-Jones and J. Beer, eds.), pp. 70–91, Guilford Press, 1st ed., 2012.

[83] H. Edmonds, L. Couture, S. Stolzy, and M. Paloheimo, "Quantitative surface electromyography in anesthesia and critical care," *International Journal of Clinical Monitoring and Computing*, vol. 3, pp. 135–145, 1986.

[84] M. Struys, L. Versichelen, E. Mortier, D. Ryckaert, J. De Mey, C. De Deyne, and G. Rolly, "Comparison of spontaneous frontal EMG, EEG power spectrum and bispectral index to monitor propofol drug effect and emergence," *Acta Anaesthesiologica Scandinavica*, vol. 42, pp. 628–636, 1998.

[85] T. Lapinlampi, H. Viertiö-Oja, M. Helin, K. Uutela, M. Särkelä, A. Vakkuri, G. Young, *et al.*, "Algorithm for quantifying frontal EMG responsiveness for sedation monitoring," *The Canadian Journal of Neurological Sciences*, vol. 41, pp. 611–619, 2014.

[86] D. Altman, *Practical statistics for medical research.* Chapman and Hall, 1st ed., 1991.